

**UNIVERZITA KARLOVA V PRAZE**

**Přírodovědecká fakulta**

Katedra sociální geografie a regionálního rozvoje

Studijní program: Geografie

Studijní obor: Učitelství geografie pro střední školy



Mgr. Andrea STRAMOVÁ

**PREZENTACE KVALITY ŽIVOTNÍHO PROSTŘEDÍ A JEHO VLIVU  
NA LIDSKÝ ORGANISMUS POMOCÍ GEOSTATISTICKÝCH NÁSTROJŮ,  
MULTIDISCIPLINÁRNÍ PŘÍSTUP**

PRESENTATION OF THE ENVIRONMENTAL QUALITY AND ITS INFLUENCE  
ON THE HUMAN ORGANISM THROUGH GEOSTATISTIC TOOLS,  
MULTIDISCIPLINARY APPROACH

Diplomová práce

Vedoucí diplomové práce: RNDr. Jakub Trubač, Ph.D.

Praha 2018

### **Prohlášení**

Prohlašuji, že jsem závěrečnou práci zpracoval/a samostatně a že jsem uvedl/a všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne 20. 6. 2018

.....

Mgr. Andrea Stramová

### **Poděkování**

Ráda bych zde poděkovala svému vedoucímu práce, RNDr. Jakubovi Trubačovi, Ph.D., za trpělivost a hodnotné rady; dále MUDr. Jiřímu Němečkovi za umožnění proniknout do problematiky štítné žlázy a pomoc při získávání endokrinologických dat; poté Přírodovědecké fakultě Univerzity Karlovy v Praze, za přístup k programům, bez kterých by práce nemohla vzniknout. Za spolupráci také děkuji společnosti ArcData, která poskytla hodnotné rady a školení při práci se Story Map. Dále bych chtěla poděkovat jednotlivým institucím (Výzkumný ústav Silva Taroucy pro krajinu a okrasné zahradnictví, Český hydrometeorologický ústav, Ředitelství silnic a dálnic, Státní zdravotní ústav, Česká geologická služba a Endokrinologický ústav) za vstřícné jednání a poskytnutí dat. Stejně tak děkuji za poskytnutí několika fotografií Jiřímí Jirouškovi. A nakonec, i když neméně, své rodině a blízkým za morální a materiální podporu v průběhu tvorby.

## **Abstrakt**

V současné době existuje velké množství dat, které charakterizují životní prostředí na základě různých parametrů. Existence takového množství datových sad je náročná na porozumění a komplexní vizualizaci. Z toho důvodu je nutné hledat nástroje pro jejich vhodné zobrazení. Cílem této práce je přehledně veřejnosti prezentovat stav životního prostředí v jejich místě bydliště a zároveň prezentovat možný vliv těchto faktorů na jejich zdraví. Aby bylo cíle dosaženo bylo nejprve nutné jednotlivé sady dat získat a připravit je, jak pro vstup do jednotlivých analýz, tak pro srozumitelnou prezentaci. V práci jsou popsány jednotlivé aspekty životního prostředí v zájmovém území Příbramsko. Následně se práce věnuje podrobnému rozboru prostorového uspořádání medicínálních dat. Součástí práce je i analýza O2PLS jež dá za vznik regresnímu modelu pro parametry funkce štítné žlázy. Výstupem práce je webová aplikace, jež srozumitelně prezentuje vstupní data (životního prostředí, medicínální), prostorové analýzy a analýzu O2PLS. Díky této práci vznikl postup přípravy a geostatistické analýzy různých typů dat a zároveň i podrobný popis prezentace těchto informací.

**Klíčová slova:** GIS, O2PLS, geomedicína, životní prostředí, story mapy, endokrinologie

## **Abstract**

Nowadays exists large amount of data which characterizes environment on base of different parametres. Due to existence of such large amount of data which is difficult to understand and comprehensively visualize, it is necessary to find tools for their proper visualization. The main goal of this thesis is present current environment in place of residence and show possible influence of environmental factors on health. For achieving the goal it was necessary to obtain individual sets of data and prepare them for each analysis and also for their presentation. In thesis there are described various aspects of environment in the area of Příbram region. Also in text is described in detail spatial determined medical data. Thesis contains O2PLS analysis which will create regression model for thyroid function parameters. Output of whole thesis is creating a web application that clearly presents input data (environmental and medical), geostatical and O2PLS analysis. Due to this work has been developed procedure how to prepare a analyze different type of data and also created detailed description of presentation results.

Till now has not been proven what aspects can affect thyroid function. Many diseases are affected by the environment. This thesis focuses on influence of environment on thyroid function. Main goal of this thesis is to propose methodology for spatial and statistical description of medical data, reveals regularity in spatial arrangement and creates a regression model. The spatial arrangement of the data is described by the method of the nearest neighbor and local Moran I. Then O2PLS method is applied and its proposed a regression model. In data was identified clustering in the vicinity of towns. In the thesis regression model which reflects aspects of the environment (concentration of As, Cd, Pb, Pb and Se in water) was proposed. For other environmental parameters has not been proven any evidence of effect on the thyroid function. This work is useful for using spation data analysis in the endocrinology field. Due to this work has been developed procedure how to process medical data and extract information in the context of spation data.

**Keywords:** GIS, O2PLS, geomedicine, environment, story map, endocrinology

## Obsah

1	Úvod.....	10
2	Předchozí výzkumy .....	12
2.1	Epidemiologie .....	12
2.2	GIS a životní prostředí.....	13
2.3	Geomedicína v ČR .....	14
3	Aplikované metody .....	17
3.1	GIS .....	17
3.1.1	ArcGIS for Desktop .....	19
3.1.2	ArcGIS online .....	28
3.2	Obousměrné ortogonální projekce do latentní struktury (O2PLS/OPLS) .....	32
4	Případová studie – Příbramsko .....	35
4.1	Popis zájmového území .....	35
4.2	Vstupní data .....	38
4.2.1	Endokrinologie .....	38
4.2.2	Endokrinologická data .....	39
4.2.3	Geoparametry .....	42
4.2.4	Data z biomonitoringu zátěže přírody a z analýz pitné vody .....	45
4.2.5	Data od ČHMÚ a intenzita dopravy.....	55
4.3	Prostorová analýza.....	62
4.3.1	Analýza prostorových vzorů .....	62
4.3.2	Variogramy endokrinologických parametrů .....	64
4.4	O2PLS/OPLS .....	75
4.4.1	O2PLS při nastavení závislé proměnné všech lékařsko diagnostických parametrů	78
4.4.2	OPLS při nastavení závislé proměnné VOLGUT .....	85
4.4.3	OPLS při nastavení závislé proměnné TSH .....	88
4.4.4	OPLS při nastavení závislé proměnné FT3.....	91
4.4.5	OPLS při nastavení závislé proměnné FT4.....	94

4.4.6	OPLS při nastavení závislé proměnné TGL .....	96
4.4.7	OPLS při nastavení závislé proměnné RAS .....	99
5	Diskuse výsledků .....	104
6	Závěr .....	108
7	Literatura .....	110
Seznam příloh.....		117

## Seznam tabulek, obrázků a grafů

Tab. 1.	Hladina jodurie v těle.....	40
Tab. 2.	Referenční hodnoty objemu štítné žlázy .....	40
Tab. 3.	Objem štítné žlázy.....	40
Tab. 4.	Hladina TSH v těle .....	41
Tab. 5.	Hladina FT3 v těle .....	42
Tab. 6.	Hladina FT4 v těle .....	42
Tab. 7.	Hodnocení distribuce prvků v mechu .....	50
Tab. 8.	Hodnocení vlivu koncentrace prvků v mechu na funkci štítné žlázy .....	51
Tab. 9.	Hodnocení vlivu pitné vody na funkci štítné žlázy .....	53
Tab. 10.	Základní ukazatele rozložení hodnot u Jodurie .....	64
Tab. 11.	Základní ukazatele rozložení hodnot u Volgut.....	66
Tab. 12.	Základní ukazatele rozložení hodnot TSH .....	68
Tab. 13.	Základní ukazatele rozložení hodnot FT3 .....	70
Tab. 14.	Základní ukazatele rozložení hodnot FT4 .....	72
Tab. 15.	Vztah mezi ukazatelem alergie a vnějšími faktory hodnocený modelem O2PLS, prediktivní komponenta .....	79
Tab. 16.	Vztah Alergie vs. Prediktory hodnocený vícenásobnou regresí.....	82
Tab. 17.	Vztah VOLGUT a vnější faktory hodnocený modelem OPLS.....	85
Tab. 18.	Vztah Volgut vs. Prediktory hodnocený vícenásobnou regresí .....	87
Tab. 19.	Vztah TSH a vnější faktory hodnocený modelem OPLS .....	89
Tab. 20.	Vztah TSH vs. Prediktory hodnocený vícenásobnou regresí.....	90
Tab. 21.	Vztah FT3 a vnější faktory hodnocený modelem OPLS.....	92
Tab. 22.	Vztah FT3 vs. Prediktory hodnocený vícenásobnou regresí .....	93
Tab. 23.	Vztah FT4 a vnější faktory hodnocený modelem OPLS.....	94
Tab. 24.	Vztah FT4 vs. Prediktory hodnocený vícenásobnou regresí .....	95
Tab. 25.	Vztah TGL a vnější faktory hodnocený modelem OPLS .....	97
Tab. 26.	Vztah TGL vs. Prediktory hodnocený vícenásobnou regresí.....	98
Tab. 27.	Vztah RAS a vnější faktory hodnocený modelem OPLS .....	100
Tab. 28.	Vztah RAS vs. Prediktory hodnocený vícenásobnou regresí.....	102
Obr. 1	Vztah mezi GIS a CAD, DBMS,CAC a DPZ .....	18
Obr. 2	Python kód pro anonymizaci zdravotních dat .....	20
Obr. 3	Normální (Gaussovo) rozdělení .....	23
Obr. 4	O2PLS modelující dvě datové tabulky.....	33
Obr. 5	Haldy na Příbramsku .....	37
Obr. 6	Štítná žláza .....	38
Obr. 7	Prezentace parametru Volgut pomocí webové aplikace .....	41
Obr. 8	Prezentace radonu v objektech pomocí webové aplikace .....	45
Obr. 9	Lokace sebraných vzorků mechu .....	46
Obr. 10	Letecký pohled na dopadovou plochu v Brdech.....	47
Obr. 11	Distribuce Pb v ČR v roce 2000 .....	48
Obr. 12	Směr větru v ČR.....	49
Obr. 13	Prezentace parametru Pb pomocí webové aplikace .....	50
Obr. 14	Letecký snímek na lesní porost v Brdech.....	52
Obr. 15	Vodní nádrž Octárna .....	55
Obr. 16	Komín teplárny Příbram.....	56
Obr. 17	Průměrná koncentrace SO <sub>2</sub> v ovzduší v letech 2007- 201.....	57
Obr. 18	Průměrná koncentrace NO <sub>2</sub> v ovzduší v letech 2007 – 2011.....	59
Obr. 19	Dálnice v České republice. ....	60



Obr. 20	Prezentace parametru NO <sub>2</sub> pomocí webové aplikace .....	60
Obr. 21	Ukázka znečištěného ovzduší .....	61
Obr. 22	Analýza prostorových vzorů.....	63
Obr. 23	Algoritmus pro analýzu výskytu nemoci s vnějšími faktory.....	77
Graf 1.	Postupná úprava hodnot pro parametr TSH.....	24
Graf 2.	Histogram naměřených hodnot Jodurie .....	64
Graf 3.	Variogram mrak naměřených hodnot Jodurie.....	65
Graf 4.	Analýza trendu u naměřených hodnot Jodurie .....	65
Graf 5.	Variogram pro naměřené hodnoty Jodurie .....	66
Graf 6.	Histogram naměřených hodnot Volgut .....	67
Graf 7.	Variogram mrak naměřených hodnot Volgut.....	67
Graf 8.	Analýza trendu u naměřených hodnot Volgut.....	67
Graf 9.	Variogram pro naměřené hodnoty Volgut .....	68
Graf 10.	Histogram naměřených hodnot TSH.....	69
Graf 11.	Variogram mrak naměřených hodnot TSH .....	69
Graf 12.	Analýza trendu u naměřených hodnot TSH.....	69
Graf 13.	Variogram pro naměřené hodnoty TSH.....	70
Graf 14.	Histogram naměřených hodnot FT3 .....	71
Graf 15.	Variogram mrak naměřených hodnot FT3.....	71
Graf 16.	Analýza trendu u naměřených hodnot FT3 .....	71
Graf 17.	Variogram pro naměřené hodnoty FT3 .....	72
Graf 18.	Histogram naměřených hodnot FT4 .....	73
Graf 19.	Variogram mrak naměřených hodnot FT4.....	73
Graf 20.	Analýza trendu u naměřených hodnot FT4 .....	74
Graf 21.	Variogram pro naměřené hodnoty FT4 .....	74
Graf 22.	Hottelingova T <sup>2</sup> statistika.....	76
Graf 23.	Komponentní váhy prediktivní komponenty vyjádřené jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	81
Graf 24.	Výsledek vícenásobné regrese pro Alergie .....	84
Graf 25.	Komponentní váhy prediktivní komponenty VOLGUT vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	86
Graf 26.	Výsledek vícenásobné regrese pro Volgut.....	88
Graf 27.	Komponentní váhy prediktivní komponenty TSH vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	89
Graf 28.	Výsledek vícenásobné regrese pro TSH .....	91
Graf 29.	Komponentní váhy prediktivní komponenty FT3 vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	92
Graf 30.	Výsledek vícenásobné regrese pro FT3.....	94
Graf 31.	Komponentní váhy prediktivní komponenty FT4 vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	95
Graf 32.	Výsledek vícenásobné regrese pro FT4.....	96
Graf 33.	Komponentní váhy prediktivní komponenty TGL vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	98
Graf 34.	Výsledek vícenásobné regrese pro TGL .....	99
Graf 35.	Komponentní váhy prediktivní komponenty RAS vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti.....	101
Graf 36.	Výsledek vícenásobné regrese pro RAS .....	103

# 1 Úvod

Tato diplomová práce aplikuje multidisciplinární přístup geostatistické a geoinformační analýzy (GIS) aplikované na geochemická data z životního prostředí a data zdravotní. Zároveň práce obsahuje podrobnou část věnovanou přehledné prezentaci těchto dat a získaných výsledků analýz pomocí webové aplikace (ArcGis Online), a to nejen pro vědeckou obec, ale i širokou laickou veřejnost, kteří mají zájem vědět jak „čisté“ je jejich okolí, kde žijí, či hodlají žít.

V roce 2000 až 2005 probíhal ve čtyřech částech České republiky sběr dat z vyšetření štítné žlázy různých věkových skupin obyvatel. Jednalo se o oblasti s rozdílnou geologickou situací a stupněm antropogenního znečištění (Příbramsko, Klatovsko, Novobystřicko a Ždárské vrchy). Získaná data byla následně zanalyzována pouze medicíně. Avšak data obsahovala další velké množství informací o vysoké podrobnosti, které se daly dále využít. Na základě prostorové informace jsem se rozhodla aplikovat geostatistické metody, které by umožnily detaily ještě více prozkoumat a objevit prostorové korelace. Podstatou geostatistických analýz je právě poloha. K dispozici byla data o poloze pacientů, ale na základě této sady dat ještě nebylo možné vytvářet podrobné závěry o vlivu životního prostředí na funkci štítné žlázy. Z toho důvodu byly získány ještě datové sady, které charakterizují znečištění ovzduší, pitné vody a celkovou zátěž přírody společně s geologickými parametry. Díky přístupu k těmto datovým sadám bylo možné naplnit cíle této diplomové práce.

Hlavním cílem diplomové práce bylo přehledně prezentovat veřejnosti stav životního prostředí v jejím místě bydliště a zároveň přiblížit možný vliv těchto faktorů na zdravotní stav. Ideálním způsobem jak veřejnost informovat, by byl vznik webové aplikace, která by přehledně ukazovala životní prostředí v celé České republice a zároveň identifikovala místa s negativním vlivem na zdraví člověka. Projektu o takovém měřítku podrobnosti. Jakého by bylo zapotřebí, nelze dosáhnout v jedné diplomové práci. Z toho důvodu bylo zájmové území v této práci zúženo na jeden okres.

Níže jsou představeny dílčí cíle diplomové práce, které vedly k naplnění cíle hlavního:

- Získat prostorová data, která v dostatečné míře popisují životní prostředí v zájmovém území.
- Data připravit pro vstup do analýz a pro prezentaci veřejnosti.
- Popsat jednotlivé aspekty životního prostředí vzhledem k jejich výskytu v zájmovém území.

- Aplikovat geostatistické metody pro popis chování zdravotních dat v prostoru.
- Navrhnout regresní model, který objasní souvislost parametrů funkce štítné žlázy s parametry životního prostředí.
- Srozumitelně podklady a výsledky práce prezentovat široké veřejnosti.
- Navrhnout postup, který bude aplikovatelný na další zájmové území, pro které jsou zdravotní data k dispozici.

Diplomová práce je rozdělena do šesti kapitol. V úvodu je uveden hlavní cíl práce. Dále následuje kapitola věnující se již provedeným výzkumům, jak v oblasti geoinformačních analýz, poznání vlivu aspektů životního prostředí na funkci štítné žlázy, tak v oblasti prezentace dat. Poté se práce zabývá aplikovanými metodami s cílem získat takový postup, který by mohl být použit i na jiné datové základny. Čtvrtá kapitola se již věnuje případové studii Příbramska. Zde jsou podrobně popsána vstupní data, kterými byla data zdravotní a parametry životního prostředí. Dále se v práci nalézá podkapitola, která je věnovaná prostorovým analýzám endokrinologických parametrů. A na jejímž závěru jsou popsány výsledky analýzy výskytu nemoci s faktory prostředí s aplikací metody O2PLS. V práci je také k dispozici diskuze dosažených výsledků se současným stavem poznání. V závěru jsou shrnuty výsledky celé práce.

## 2 Předchozí výzkumy

Existuje celá řada možností, jakými lze popsat životní prostředí či populaci, zvířata a rostliny. Díky tomu existuje velké množství dat, které dané téma popisují vždy z pohledu rozdílného parametru. Zároveň existují disciplíny, které taková data shromažďují a následně je zkoumají (epidemiologie, regionální geografie a další). V současné době ale neexistují jednotná databáze všech aspektů životního prostředí, která by umožňovala komplexní analýzy a zhodnocení. Z toho důvodu jsou nejčastěji vytvářeny studie, které hodnotí vliv pouze jednoho faktoru životního prostředí na jeden zdravotní parametr.

Diplomová práce předkládá data jak geomedicinální, tak i geochemická v co největší možné šíři. Z toho důvodu zde byl nutný multidisciplinární přístup k samostatným datům, ale i k jejich zpracování a následné prezentaci. Informace, které data nesou, jsme schopni číst pomocí geostatistických metod.

### 2.1 Epidemiologie

Jak již bylo řečeno v úvodu, předkládána diplomová práce spojuje velké množství vědních disciplín v jeden celek. Jednou z vědních disciplín je epidemiologie. Epidemiologie představuje vědní obor, který se podle Světové zdravotnické organizace zabývá studiem rozdělení a příčin nemocí a událostí spjatých se zdravotním stavem lidské populace s následnou aplikací zjištěných poznatků při řešení zdravotních problémů (Bencko 2003). Další vědní disciplínou, která významně zasahuje do této práce je statistika. Dle slovníku cizích slov je statistika vědním oborem zabývajícím se sběrem, analýzou a zpracováním informací, které kvantitativně charakterizují zákonitosti jevů ve spojitosti s jejich kvalitativním obsahem (SCS 2018). Data, která byla v práci zpracovávána, jsou vícerozměrná, a z toho důvodu byly aplikovány metody vícerozměrné analýzy, které umožňují propojení epidemiologie se statistikou.

Ještě předtím, než ale byly známy složité počítačové vícerozměrné analýzy, odhalil John Snow souvislost mezi znečištěním pitné vody a nákazou cholery, a to již v roce 1854. Jeho postup byl sice velice jednoduchý, ale účinný a dokázal správně označit zdroj nákazy. Díky tomuto činu je nazýván otcem epidemiologie (BBC 2014). S postupem času se vědci věnovali stále novějším tématům.

#### **Případová studie Černobyl**

Velice významným tématem, které úzce souvisí s funkcí štítné žlázy, byla Černobylská havárie z 26. dubna 1986, kdy byla prováděna bezpečnostní zkouška, která skončila dvěma výbuchy. Tyto výbuchy způsobily postupné uvolňování vzácných plynů a těžkých štěpných

produktů (Xe, Kr, J, Te a Cs) do okolního prostředí. Důsledkem byl vznik radioaktivního oblaku, který postupně vlivem větru přecházel přes Evropské státy, konkrétně Švédsko, Finsko, následně přes střední a západní Evropu (Drábová a kol. 2001). První signály úniků byly zaznamenány Švédskem a to den po výbuchu 27. 4. 1986 (Tomanová 2005).

Československem postupně prošly tři radioaktivní vzdušné masy, kdy hodnoty radioaktivity dosahovaly až  $257 \text{ Bq/m}^3$ . V Praze - Libuši byla naměřena 30. 4. 1986 objemová aktivita  $^{131}\text{I}$   $70 \text{ Bq/m}^3$  (IHE 1987). Následně byly prováděny testy půdy, vody, potravin a zemědělských produktů z důvodu prevence možného vystavení lidského organismu radioaktivitě (IHE 1987). Nebylo však možné této katastrofě plně zamezit. Byly zaznamenány zvýšené počty potratů plodů v 8. - 12. týdnu, kdy převažovaly potraty mužského plodu (Peterka a Peterková 2004). Hlavně u dětí byly identifikovány zvýšené počty rakoviny štítné žlázy kvůli spadu radioaktivního jódu v prvním roce po Černobylské havárii (Knechtsberger [online]). V oblastech Bryansk, Kyjev a Zhitor byl dokládán výskyt rakoviny štítné žlázy u 35,6 % dětí (Ashizawa a kol. 1997). Taktéž ve Velké Británii byl proveden výzkum vlivu radioaktivního jódu na výskyt rakoviny štítné žlázy (Hill a kol. 1986). I v současné době je Černobylská havárie velký unikát, který stále vědce zajímá. Z toho důvodu v roce 2013 vznikl článek od Kurttio a kol., který se věnuje souvislosti Černobylského spadu radioaktivního materiálu s výskytem rakoviny ve Finsku. Práce se snaží zpětně dokázat korelaci mezi radioaktivním jódem a rakovinou štítné žlázy. Autoři měli k dispozici podrobná data z národních statistik Finska jak zdravotnická, tak z měření radioaktivity. Tato práce vyloučila vliv radioaktivity na rakovinu prsu, plic a prostaty. Naopak zajímavých výsledků bylo dosaženo při korelaci s rakovinou štítné žlázy u mladých mužů, kde byla prokázána statistická významnost. Obdobnému tématu se věnovala švédská studie Tondela a kol. (2004), jejímž výsledkem bylo pouze opatrné prohlášení, že havárie mohla mít vliv na zvýšený nárůst rakoviny na území Švédska. Na území Běloruska a později i na Ukrajině a v Ruské federaci se rovněž objevovali zprávy o zvýšeném počtu rakoviny štítné žlázy u dětské populace. U dospělé populace přesvědčivé důkazy o zvýšeném výskytu rakoviny štítné žlázy neexistují (Reiners 2000).

## 2.2 GIS a životní prostředí

Tato diplomová práce se zabývá mimo jiné vlivem životního prostředí na funkci štítné žlázy, případně na lidské zdraví. Jelikož pojem životní prostředí v sobě nese významný prostorový kontext, bylo zapotřebí do práce zařadit další vědní disciplínu, a to geografické informační systémy (GIS). Díky využití metod této disciplíny je možné data sbírat, upravovat, analyzovat a následně také prezentovat. Užitečnost GIS při řešení epidemiologických problémů již byla

několikrát prokázána (Elliott a kol. 2001, Grundloch 2009, , Wang a kol. 2015) a s postupem vědy a vývojem počítačových technologií bude mít jistě v budoucnu významný přínos při řešení složitějších problémů. V článku Bùlvikena a kol. (1997) byly nástroje GIS použity k prokázání korelace mezi mírou mortality a roztroušenou sklerózou s koncentrací atmosférických spálenin a Se v ovzduší a také s koncentrací Ca a K v nadbytečném sedimentu. Ve studii zabývající se rizikem života populací v okolí skládek (Elliott a kol. 2001) byly vhodně nástroje GIS využity pro zpracování dat i pro následné analýzy. Pro mapování a identifikaci zdrojů škodlivých látek nebezpečných lidskému organismu byly využity nástroje GIS v práci Zhanga (2006).

Práce Poulstrupa (2004) se věnovala třem oblastem, které se lišili stupněm znečištění ovzduší dioxiny v souvislosti s výskytem rakoviny. Podobná práce zaměřující se na znečištění ovzduší se věnovala pouze jednomu městu - Jinan v Číně. Zde se Wang a kol. (2015) pokusili pomocí nástrojů GIS zjistit korelaci mezi znečištěním ovzduší a výskytem bronchitidy u konkrétních jedinců, ke kterým měli k dispozici zdravotní záznamy i záznamy místa bydliště a pracoviště. V práci Brodyové a kol. (2004) byly GIS použity v case-control studii, která hledala možnou souvislost vlivu pesticidů na rakovinu prsu u žen na území státu Massachusetts. V článku Shahida a kol. (2014) je korelováno znečištění vody a závažná onemocnění žaludku rovněž pomocí GIS nástrojů. V článku Henke a Petropoulos (2013) byla úspěšně prokázána korelace vzdálenosti rekreačních oblastí a délky dožití na území Velké Británie.

Z daného výčtu článků je patrné, že nezáleží na poloze zkoumaného místa na Zemi. Podstatnou částí jsou zde podkladová data. Jednotlivé články se od sebe liší právě podrobností dat. V několika člancích se pracuje s jednotlivci, kteří již byli nakaženi nebo s náhodným výběrem (Wang a kol. 2015, Elliott a kol. 2001). Data pro tyto články byla velmi podrobná a zaznamenávala jak místo bydliště, tak místo pracoviště a množství zdravotních parametrů. V jiných člancích se naopak pracuje s daty za celé obce nebo městské části (Kurtio a kol. 2013, Shahida a kol. 2014). V člancích byl často využíván vždy jen jeden parametr životního prostředí. Například se jednalo o vzdálenost od skládek, stupeň znečištění ovzduší dioxiny a další. Propojení více aspektů životního prostředí v jedné studii je náročné. Prvním aspektem je získání potřebných datových sad, které by popisovaly životní prostředí na základě různých parametrů. Druhým faktorem je nutný multidisciplinární přístup, tedy propojení více vědních disciplín aby bylo možné proniknout do problematiky a následně interpretovat závěry.

## **2.3 Geomedicína v ČR**

Metody prostorové analýzy se nepoužívají pouze v zahraničí, ale byly již několikrát využity také v ČR. Nejprve byly geoinformační systémy ve vztahu k životnímu prostředí využívány spíše

pro zpracování určitých dat (Modráčková 1999) nebo pro modelování určité zátěže na obyvatelstvo (Karlova 2001, Jeřábková 2004). Významný krok kupředu vznikl propojením Endokrinologického ústavu s odborníky z Univerzity Karlovy v Praze. Endokrinologický ústav poskytl pro některé práce velice podrobnou datovou základnu týkající se zdravotního stavu pacienta se zaměřením na výsledky vyšetření štítné žlázy. Díky tomuto propojení vznikla v roce 2005 diplomová práce Tomáše Görnera a disertační práce Renáty Tomanové a následně práce Grundlocha (2009) a Stramové (2017).

Práce Görnera se věnuje prokázání vlivu možného vlivu radioaktivního záření geologického podloží hornin a antropogenní radioaktivity na hodnotu jodurie u dětí a ne karcinom plic u dospělé populace. Pro své modelové území si Görner zvolil Klatovy. V práci jsou podrobně rozebrány zdroje radioaktivního záření a jeho možnost ovlivnění lidského zdraví. Následně jsou zde popsány zdroje dat, které se převážně věnují zdrojům radioaktivního záření a jejich naměřeným výsledkům, jedná se tedy pouze o jeden aspekt životního prostředí. Z medicínálního hlediska se práce věnuje u štítné žlázy pouze jednomu ukazateli, a to jodurii, a u rakoviny plic počtu tumorů. GIS nástroje zde byly použity převážně pro zpracování dat a jejich prezentaci pomocí mapových výstupů. Veškeré statistické analýzy byly prováděny pomocí jednoduchých tabulek, kdy mezi sebou byly porovnávány skupiny obyvatel podle toho, na jakém místě žijí. Práce nenalezla žádný vliv geologického podloží ani antropogenní radioaktivity na zkoumané lékařské diagnózy. Nicméně jedná se o jednu z prvních prací, která dávala do souvislosti mezioborové podklady a informace.

Rigorózní práce Tomanové (2005) se věnuje vlivu geologie a radioaktivního záření pouze na výskyt dětské thyreopatie na testované lokalitě Příbramsko. Práce je velice kvalitní z pohledu informací o geologii a radioaktivním záření. Data charakterizující životní prostředí jsou zaměřena hlavně na geologické podloží, které ale nebylo v dřívějších pracích bráno v potaz. Příprava dat pro vstup do korelačních analýz je poměrně nepřehledná a chybí zde její konkrétní popis. Tato práce stejně jako práce Grundlocha (2009) čerpala informace o prostředí pouze z polohy bydliště vyšetřeného pacienta. V práci Stramové (2017) již byly informace čerpány i z polohy školy, kterou pacient pravděpodobně navštěvovala na základě vymezení spádových oblastí jednotlivých škol. Tomanová v práci použila Spearmanův koeficient pořadové korelace. Celý datový soubor byl rozdělen na základě několika parametrů do různých skupin, tím ale bohužel došlo ke značné ztrátě prostorové informace. Podstatným jevem, který byl v práci zkoumán, byla hodnota jodurie, což je parametr, díky němuž můžeme zjistit funkci štítné žlázy, ale existují i podstatnější parametry např. Volgut, TSH, FT3 a případně FT4.

V práci nebyly nalezeny žádné statistické významné výsledky, ale byla odhalena velmi slabá závislost mezi jodurí a koncentrací uranu

Obdobnému tématu i s podobným postupem se věnoval i Grundloch (2009), ale pro změnu v lokalitě Krucemburku, Radostína nad Oslavou, Dolních Rožinek a Budišova. V této práci také došlo k markantnímu zvětšení datové základny, která zahrnovala už i data o znečištění ovzduší či vod společně s daty o geologickém podloží. Celkově byla práce metodicky rozšířena o využití nástrojů GIS pro přípravu dat. Zde se naopak pomocí Spearmanova koeficientu pořadové korelace podařilo prokázat slabé korelace mezi environmentálními aspekty (koncentrace antimonu, železa a arsenu v pitné vodě) a parametry štítné žlázy.

Poslední prací věnující se této tématice byla práce Stramové (2017). Tato práce již podrobně rozebírala přípravu jednotlivých datových vrstev. Zároveň se práce nesoustředila pouze na jeden ukazatel funkce štítné žlázy. V práci byly ve větší míře využity nástroje GIS pro prostorové analýzy, u nichž bylo výsledkem konstatování shlukového rozdělení výskytu nemocných pacientů. Posun vpřed byl také v metodice korelací. Byla zde aplikována metoda O2PLS, která vedla k vytvoření jednoduchého regresního modelu. Výsledkem bylo vysvětlení vlivu několika parametrů životního prostředí (As, Pl, Pb, Cd, a Se v pitné vodě) na funkci štítné žlázy.

Všechny tyto práce byly jakýmsi předchůdci této diplomové práce. Na základě analýz Stramové (2017) byla zde rozšířena datová základna parametrů pitné vody, která byla prokázána jako významná při ovlivňování funkce štítné žlázy. Jelikož se předchozí práce věnovaly primárně prokázání vlivu jednotlivých složek životního prostředí na funkci štítné žlázy, nebyl zde prostor pro popis těchto složek. V této práci je popisu složek vyhrazen prostor. Díky osvědčení jednotlivých metod prostorových analýz, zde byla možnost v analýzách pokračovat a zhodnotit zdravotnická data na základě jejich variogramu a tím popsat jejich chování v prostoru. Další možností rozšíření bylo aplikovat metodu O2PLS/OPLS, která se osvědčila. Bohužel výstupem zmíněných prací bylo několik málo mapových výstupů a tabulek, které nebyly prezentovány pro veřejnost. Proto se tato diplomová práce věnuje i veřejné prezentaci veškerých datových podkladů a výstupů z jednotlivých analýz.

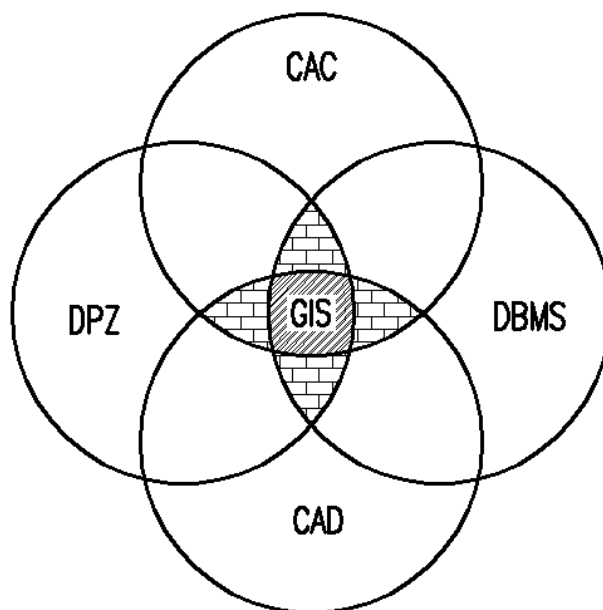


### 3 Aplikované metody

V následující části jsou představeny jednotlivé metody, které sloužily k naplnění cílů diplomové práce. Kapitola je rozdělena do dvou podkapitol. První se věnuje problematice GIS a jejího využití. Nejprve je popsán proces přípravy dat jak pro vstup do prostorových analýz, tak pro vstup do webové aplikace. Představeny jsou zde aplikované metody prostorové analýzy dat. V závěrečné části podkapitoly je popsán postup tvorby webové aplikace pro prezentaci dat veřejnosti. Druhá podkapitola se věnuje analýze výskytu nemoci s faktory životního prostředí, a to za použití metody O2PLS. Zde je podrobně popsán teoretický rámec této metody společně s dalšími aplikovanými metodami, které jsou pro O2PLS podstatné.

#### 3.1 GIS

Rozvoj GIS je v posledních letech mimořádný a díky své šíři záběru se dostávají do mnoha vědních oborů. Z těchto důvodů je důležité položit si otázku: „Co jsou geoinformační systémy?“ V současné době se používá název geoinformační systémy, GIS nebo geografické informační systémy. V literatuře se využívají všechny názvy a existuje také velké množství definic. Tollingerová (1996) uvádí, že GIS umožňuje pořizovat, využívat a aktualizovat rozsáhlé databáze s tematicky různorodými prostorovými daty. Dále dovoluje analyzovat a kombinovat údaje obsažené v databázi, a tím určovat polohu, kvalitu, stav, vývoj nebo možné důsledky nejrůznějších situací. Voženílek (1998) definuje GIS v geografii jako organizovaný, počítačově založený systém hardwaru, softwaru a geografických informací vyvinutý ke vstupu, správě, analytickému zpracování a prezentaci prostorových dat s důrazem na jejich prostorové analýzy. Kolář (2003) definuje GIS jako propojení počítačového návrhářství (CAD), databázových systémů (DBMS), počítačové kartografie (CAC) a dálkového průzkumu Země (DPZ), jak je vidět na Obr. 1 .



**Obr. 1 Vztah mezi GIS a CAD, DBMS, CAC a DPZ**  
(Zdroj: Kolář 2003)

Štych a kol. (2008) tvrdí, že GIS lze chápat jako informační systém, který uchovává zejména polohově vztažená data a umožňuje jejich sběr, kontrolu, organizaci, aktualizaci, distribuci a vizualizaci. Důležitým aspektem systému je možnost vykonávat analýzy prostorových dat. Společnost, zabývající se vývojem softwaru ArcGIS ESRI (2018), definuje GIS jako systém, který nám umožňuje vizualizovat, dotazovat se, analyzovat a interpretovat data a zároveň porozumět vztahům, vzorům a trendům.

V současné době je GIS chápán jako ucelený systém jednotlivých komponentů (Štych a kol. 2008)

- Technologie- software a hardware
- Databáze - geografické a atributové údaje
- Infrastruktura - uživatelé a podpůrné složky

GIS pracují s údaji, které jsou polohově vázány k povrchu Země. Dále dokáží řešit problémy, které kombinují dotazy na vlastnosti objektů s dotazy na jejich polohu nebo jejich vztah k dalším objektům. Výsledky GIS interpretuje pomocí přehledné mapy, ale zároveň i pomocí tabulek a textů.

Uživatelé GIS můžeme rozdělit do dvou skupin. První skupinou jsou operátoři, kteří data zpracovávají pro koncového uživatele. Operátorem je například kartograf nebo odborný referent. Druhou skupinou jsou koncoví uživatelé. Tito uživatelé se s podporou GIS rozhodují. Gis se využívá v archeologii, armádě, dopravě, na katastru nemovitostí a územního plánování,

v lesním hospodářství, při ochraně životního prostředí, na policii nebo u záchranných služeb. Tento seznam není konečný. Najde se mnoho oborů, kde je možné GIS využívat, a již dlouhou dobu se nejedná pouze o geografii.

Pro účely diplomové práce byly zvoleny programy ArcGIS for Desktop a ArcGIS online. Oba tyto programy jsou od společnosti ESRI. Společnost ESRI se dlouhodobě věnuje vývoji softwaru, který je určen pro práci s geografickými informačními systémy. Sama společnost uvádí v překladu heslo „Stavíme ArcGIS, nejsilnější software pro mapování a prostorovou analýzu na světě.“ (ESRI 2018).

### **3.1.1 ArcGIS for Desktop**

ArcGIS for Desktop je jeden z mnoha programů, které vytvořila společnost ESRI. Tento program je určen ke sběru dat a jejich správě, vytváření profesionálních map, provádění základních i pokročilých prostorových analýz a v neposlední řadě pro sdílení dat. V této diplomové práci byla využita centrální aplikace ArcMAP verze 10.3.1. Součástí programu ArcGIS for Desktop jsou ještě aplikace ArcCatalog, ArcScene a ArcAdministrator.

#### **3.1.1.1 Příprava dat**

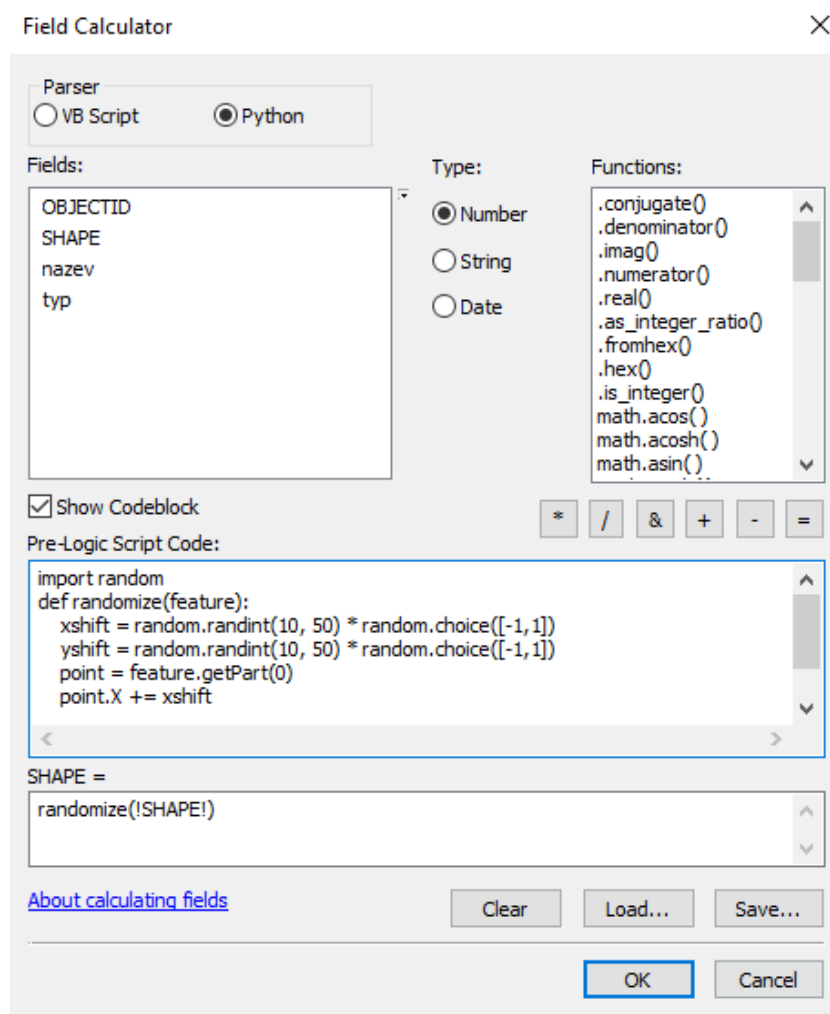
Data, která byla pro tuto diplomovou práci k dispozici, byla často obdržena v rozdílných formátech, které nebylo možné bez přípravy použít pro prezentaci v aplikaci a vytvoření prostorových analýz. Z toho důvodu bylo zapotřebí nejprve data upravit.

##### **Příprava dat pro webovou aplikaci**

Geoparametry (geologické podloží, zlomy, radiometrie) byly poskytnuty Českou geologickou službou přímo pro zájmové území Příbramsko. Zde nebylo nutné žádných úprav, jelikož data byla dodána ve formátu .shp. Posledním prezentovaným geoparametrem je obsah radonu v objektech. Tato data jsou volně dostupná na stránkách Radonového programu pro ČR. Zde je seznam jednotlivých obcí dělený podle okresů a k nim zpracované naměřené hodnoty. Tato data jsou v souboru pdf. Soubor byl nejprve přeuložen do formátu .xls. Takto připravený soubor byl připojen k vrstvě obce.shp pomocí funkce Join Table na základě názvu obce a se zachováním všech záznamů. Vznikla tak vrstva .shp s informacemi o názvu obce, počtu změřených objektů v obci, podílu objektů nad 300 Bq/m<sup>3</sup> [%], podílu objektů nad 1000 Bq/m<sup>3</sup>, podílu objektů nad 300 Bq/m<sup>3</sup> [%] a geometrickém průměru [Bq/m<sup>3</sup>].

Zdravotní parametry již byly předem zpracované v diplomové práci Stramové (2017). To znamená, že již byl k dispozici soubor ve formátu .shp, který zaznamenává místo bydliště

vyšetřených pacientů a k nim přiřazené zdravotní parametry. V diplomové práci Stramové (2017) ale nebyla řešena anonymizace dat. Tato diplomová práce však anonymizaci potřebovala, jelikož data byla prezentována ve větším rozlišení včetně podrobnosti. Pro zachování anonymity vyšetřených pacientů byly navrženy dvě možnosti řešení. První možností bylo omezení měřítka mapy s cílem, aby se uživatel nemohl přiblížit příliš blízko v mapové kompozici. Tato možnost by se dala v aplikaci vyřešit vytvořením nové podkladové mapové vrstvy, ale ve zdrojovém kódu by zůstaly zachovány GPS souřadnice, které by následně nesplnily podmínku anonymizace. Z toho důvodu byla zvolena možnost změny souřadnic, aby nevyjadřovaly skutečnou polohu adresy pacientů. Tento krok byl vytvořen v programu ArcMAP. U vrstvy zdravotních pacientů byla otevřena atributová tabulka. Následně u sloupce znázorňující polohu pacientů byla zvolena funkce Field Calculator a pomocí Python kódu byly jednotlivé souřadnice náhodně přesunuty o 10 až 50 m náhodným směrem. Tento kód a veškeré nastavení pro anonymizaci zdravotních dat je znázorněno na Obr. 2. Takto upravená data již byla připravena pro prezentaci ve webové aplikaci.



**Obr. 2** Python kód pro anonymizaci zdravotních dat

V diplomové práci byla rovněž využita vrstva zachycující základní školy na území okresu Příbram. Tato bodová vrstva byla ve vhodném formátu .shp. Pro názornost bylo pro jednotlivé školy vymezeno spádové území a až poté byla vrstva připravena pro prezentaci ve webové aplikaci.

Dalšími prezentovanými prvky ve webové aplikaci jsou parametry kvality ovzduší. Tato data jsou trojího typu, a to bodová, polygonová a liniová. Bodová vrstva znázorňuje data naměřená přímo na měřících stanicích ČHMÚ a polygonová data jsou průměrné dosažené hodnoty v celém okrese v letech 2007 až 2011. Liniová vrstva byla dodána Ředitelstvím silnic a dálnic. Znázorňují intenzitu dopravy na významných silnicích okresu Příbram. Tato data byla k dispozici již ve vhodném formátu, a proto nebylo nutné je více upravovat.

Sekce parametry biomonitoringu obsahují dvě sady dat od rozdílných subjektů. První sada dat znázorňuje koncentraci prvků v mechu a byla získána od Výzkumného ústavu Silva Taroucy pro krajinu a okrasné zahradnictví (VÚKOZ). Tato sada dat již rovněž byla pro účely této diplomové práce ve vhodném formátu získána z diplomové práce Stramové (2017), kde je podrobně popsán postup přípravy těchto dat. Druhou prezentovanou sadou dat jsou rozborů vzorků pitné vody od Státního zdravotního ústavu (SZÚ). Data byla poskytnuta za období od 1. 1. 2004 do 1. 12. 2005, a to pro 30 vodovodů (40 obcí) v okrese Příbram, které zásobují přibližně 80 % vyšetřených pacientů. Tato data byla rozdělena do samostatných .xls souborů za jednotlivé vodovody. Bylo tedy potřeba seřadit data do jedné tabulky a k vodovodům přiřadit obce, které zásobují. Tento soubor byl uložen ve formátu.xls. Takto vytvořený soubor byl následně v ArcMAP připojen funkcí Join Table k vrstvě obcí na základě názvu obce a byly zachovány všechny záznamy. Vytvořená vrstva byla uložena jako obce\_szu.shp.

Takto bylo vytvořeno 9 mapových vrstev. Mapová vrstva vytvořená v ArcMAP je uložena jako 8 záznamů, které musí být vždy v jedné složce v PC, aby vrstva fungovala a bylo možné ji zobrazit. Pro použití v ArcGIS online bylo potřeba nahrát tyto složky mapové vrstvy jako jeden soubor. K tomu byl využit free software 7-zip, díky kterému bylo možné vytvořit vždy ke každé mapové vrstvě jeden soubor formátu .zip, který obsahoval veškeré složky mapové vrstvy. Takto připravené mapové vrstvy byly nahrány do aplikace ArcGIS online.

### **Příprava dat pro analýzu O2PLS**

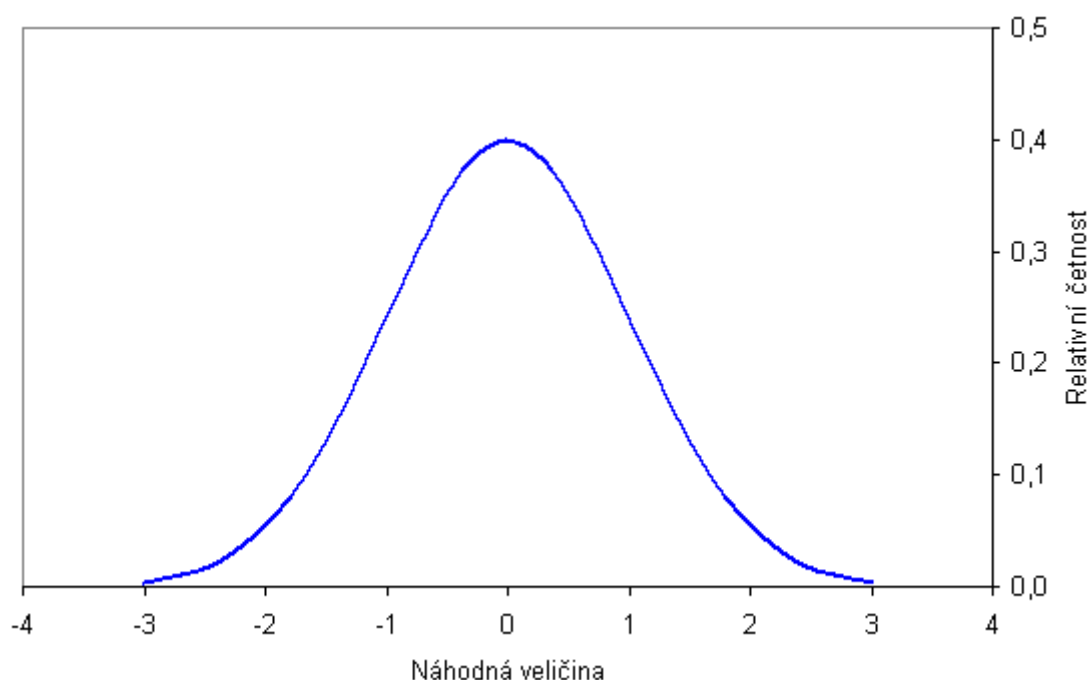
Součástí této práce je analýza s názvem obousměrné ortogonální projekce do latentní struktury (O2PLS), která je popsána v kapitole níže. Avšak pro vstup do této analýzy je nutné připravit vhodně data. Doposud bylo vytvořeno 8 samostatných vrstev, které znázorňují vybrané parametry životního prostředí a zdravotní parametry. Takto připravená data

ale nemohou vstoupit do aplikace, jelikož většina dat životního prostředí je bodového charakteru a zaznamenává hodnoty prostředí pouze v daném bodě. V diplomové práci je ale potřebné mít přehled o celé ploše. Řešením tohoto problému je interpolace prostorových dat. Interpolace na základě konkrétního nastavení statisticky vypočítá rozložení dat. Pro účely diplomové práce byla volena metoda IDW, která pracuje na předpokladu, že blízké body jsou si podobnější než body vzdálenější. Celý postup interpolace jednotlivých dat je podrobně popsán v diplomové práci Stramové (2017).

Veškeré vytvořené vrstvy byly nahrány do programu ArcMAP a připojeny k vrstvě zdravotních parametrů funkcí Join data na základě polohy jednotlivých bodů. Zachovány byly vždy veškeré záznamy. Vrstvy zlomů a intenzity dopravy navíc obsahovaly informaci ovzdálenosti pacienta ke zlomu nebo silnici. Takto vznikla nová vrstva obsahující jak zdravotní parametry, tak parametry životního prostředí.

Tato vrstva byla uložena jako tabulka formátu .xls a pro přehlednost byla tabulka rozdělena do skupin: identifikační data, binární data, kategoriální nebinární data, ordinální data a metrická data. Metrická data jsou naměřené zdravotní hodnoty a také naměřené hodnoty vnějšího prostředí pacienta. Pro následnou analýzu je nutné, aby metrická data splňovala předpoklad normality a konstantního rozptylu (homoscedasticita), tedy aby měla v ideálním případě normální (Gaussovo) rozdělení nebo alespoň rozdělení symetrické. Normální (Gaussovo) rozdělení je rozdělení pravděpodobnosti spojitě náhodné veličiny. Má tvar křivky znázorněné na Obr. 3 Ta je definována Gaussovou funkcí (vzorec 1). V dané funkci je důležitý průměr a směrodatná odchylka nebo rozptyl (Tomsa 2011).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1)$$



**Obr. 3 Normální (Gaussovo) rozdělení**  
(Zdroj: Tomsa 2011)

Normalita byla otestována v programu IBM SPSS Statistics. Naprostá většina parametrů nesplnila předpoklad normality. Z toho důvodu bylo nutno použít k jejich úpravě mocninné transformace, kterou bude primárně dosaženo hodnot šikmosti blízké nule a sekundárně i hodnoty špičatosti blízké 3 (méně důležité). Mocninná transformace je metoda primárně využívaná pro zlepšení rozdělení dat. Vhodnou transformací dat je možné stabilizovat rozptyl, přiblížit šikmost rozdělení k nule a tvar k normálnímu rozdělení (Meloun, Hill a Militký 2000). Cílem je nalézt vhodnou mocninu na základě znalostí o výběru dat. Nejzákladnější je prostá mocninná transformace (vzorec 2 a 3):

$$\text{pro } P \neq 0 \quad hp(x) = \text{sign}(x) * \text{abs}(x)^P \quad (2)$$

$$\text{pro } P = 0 \quad hp(x) = \ln(x) \quad (3)$$

Kde **abs (x)** je absolutní hodnota a **sign (x)** je znaménková funkce:

$$\text{Sign}(x) = 1 \text{ pro } x > 0, \text{ sign}(x) = -1 \text{ pro } x < 0, \text{ sign}(x) = 0 \text{ pro } x = 0.$$

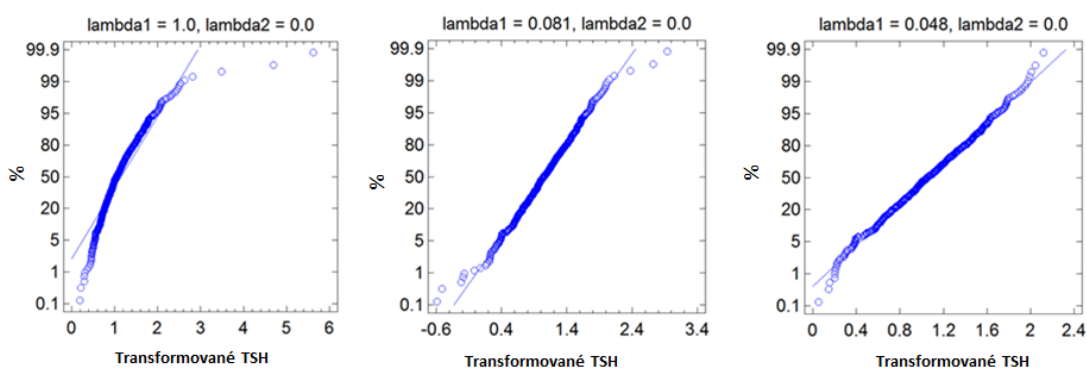
Tato transformace zachovává pořadí dat výběru a dochází k symetrizaci výběru (Meloun a kol. 2000).

Data byla nejprve posunuta přičtením konstanty ke kladným hodnotám blízkým nule. V dalším kroku byl výsledek podělen mediánem (díky tomuto pak transformace úspěšně fungovala i při vysokých mocninách). Původní soubor dat byl uložen do záložky s názvem data,

poté byla vytvořena nová záložka, která pomocí funkce KDYŽ upravovala vybraná data (jednalo se o data metrická) a automaticky je ukládala do záložky datac. Funkce KDYŽ byla nastavena takto:

KDYŽ(data!BI2="";"";KDYŽ(datap!BI\$7=1;(data!BI2+1.1\*datap!BI\$6\*(ABS(datap!BI\$11)+0.05\*ABS(datap!BI\$13)))/(datap!BI\$13+1.1\*datap!BI\$6\*ABS(datap!BI\$11));data!BI2+1.1\*datap!BI\$6\*(ABS(datap!BI\$11)+0.05\*ABS(datap!BI\$13)))).

Vzorec čerpá hodnoty z původního souboru dat a ze záložky s názvem datap. Do této záložky se automaticky vždy počítalo minimum, maximum, medián, šikmost, špičatost a byl zaznamenán i počet hodnot. Většina hodnot byla vypočítána bez problémů automaticky na základě nastavení. Následně byl formulář datac nahrán do programu Statgraphics. Zde byla pomocí funkce Describe - Numeric data - Power transformation získána hodnota  $\lambda$  (mocnina). Hodnotu  $\lambda$  bylo možné použít až po kontrole koeficientu šikmosti a z - skóre pro šikmost, které mají být blízké nule. Zároveň byl při postupu vizuálně zkontrolován kvantil-kvantilový graf pro normální rozdělení, kdy bylo někdy využito vyjmutí odlehlých hodnot. Postup úpravy grafu je znázorněn na Graf 1. Na levé části grafu jsou znázorněna data bez transformace. Na druhém grafu zleva jsou zakreslena data po automatické transformaci. V tomto případě bylo nutné odstranit odlehlé hodnoty. Na posledním grafu zleva jsou znázorněny hodnoty po automatické transformaci po vyjmutí odlehlých bodů. Po kontrole vhodnosti všech parametrů bylo možné vzít hodnotu  $\lambda$  a zkopírovat ji do původního formuláře datap v Microsoft Excel.



**Graf 1. Postupná úprava hodnot pro parametr TSH**

U medicínálních dat nastala v několika případech situace, kdy program nedokázal určit nejvhodnější hodnotu  $\lambda$ . Proto byla hodnota  $\lambda$  odhadnuta metodou půlení intervalu. Kritériem přitom byla blízkost hodnoty šikmosti a jejího z - skóre nule.



U parametrů vod bylo potřeba rozhodnout, zda je transformace nezbytná, neboť hodnoty se pohybovaly jen v malém počtu úrovní (jsou naměřeny dvě nebo tři hodnoty, které se opakují). U parametrů mechů byla ve většině případů použita transformace. Díky užití interpolaci jsou data rozmístěna pravidelně. Na grafech bylo patrné, že jsou data v území různorodá. Tuto různorodost a výkyvy bylo třeba zachovat. U dat z ČHMÚ byly prováděny obdobné kroky jako dat mechů. Na grafech dat imisního znečištění byla vidět největší soudružnost a bez úprav se použila hodnota  $\lambda$  vypočítaná programem. Tento případ nastal i u dávkového příkonu gama záření hornin. Hodnoty znázorňující průměr hodnot naměřených v místě bydliště a u škol nebylo třeba upravovat a transformace proběhly bez problémů.

Získané hodnoty  $\lambda$  byly zadány do sešitu .xls pro transformace, záložky datap a pole s názvem síla. V záložce datat pak byly vypočítány transformované hodnoty za pomoci funkce KDYZ nastavenou takto:

$$\text{KDYŽ}(\text{data!BI2}="" ; "" ; \text{KDYŽ}(\text{datap!BI\$2}=0 ; \text{LOG}(\text{datac!BI2}) ; \text{SIGN}(\text{datap!BI\$2}) * (\text{datac!BI2}^{\text{datap!BI\$2}}))$$

Posledním krokem byly retransformace získaných hodnot rekuretním vzorcem pro kontrolu. Transformovaná data pak byla dále použita pro vstup do analýz vyžadujících Gaussovske rozdělení.

### **3.1.1.2 Metody prostorové analýzy**

Metody prostorové analýzy využívají soubor technik, jež vznikly v různých oborech. Společným cílem těchto technik je analýza dat s důrazem na prostorové vztahy. Prostorové analýzy v dnešní době zaujímají významné postavení v celé řadě vědeckých oborů, v nichž byla rozpoznána jejich důležitá role k uchopení a analýze množství prostorových dat. Hlavním zájmem je nalezení nových vztahů mezi uspořádání prvků a jejich atributů ve zkoumané oblasti a zároveň modelování těchto vztahů. Hlavním cílem těchto metod je dosáhnout lepšího porozumění a předpovídání daných vztahů a uspořádání (Horák a kol. 2000). Pro účel těchto analýz rozumíme jednotlivými prvky geografické objekty a jiné objekty s prostorovou lokací, která jsou fyzické či abstraktní povahy. Častým vstupem do analýz jsou události a jevy. Na rozdíl od jiných forem analýz jsou zde podstatná atributová data a k nim lokalizace objektů (Horák 2002). Díky dostupnosti kvalitních prostorových dat a dále rozšíření použití GIS a dalšího softwaru (SPSS, Statgraphics, Simca 12+) vhodného pro práci s tímto typem dat v nedávné době vzrostl zájem o prostorové analýzy dat (Spurná 2006). Podle Fortheringham, Roberson (1994) mají prostorové analýzy dat tři hlavní cíle:

- popis událostí nebo objektů v geografickém prostoru (včetně popisu uspořádání)
- interpretace procesů, které vedly k pozorovanému stavu uspořádání objektů nebo událostí ve sledovaném prostoru (systematický průzkum)
- zlepšení schopnosti předpovídat a kontrolovat objekty nebo události v geografickém prostoru.

Všechny tyto cíle prostorové analýzy dat naplňuje i předkládaná diplomová práce a v souvislosti s ní vzniklá webová aplikace. Předkládané metody prostorové analýzy jsou vypracovány na datové sadě zdravotních parametrů, která je shodná s datovou sadou využitou v diplomové práci Stramové (2017). Jelikož část věnovaná analýze prostorových vzorů byla v diplomové práci Stramové (2017) obsáhlá a kvalitně zpracovaná, nebylo nutné pro účely této diplomové práce tuto část opakovat nebo aplikovat jiné metody. Z toho důvodu je této problematice v této práci věnována pouze malá část zabývající se metodou variogramu. Ve webové aplikaci byly uplatněny výsledky metod z diplomové práce Stramové (2017). Tyto metody jsou zde v krátkosti představeny a podrobnější popis je uveden ve zmiňované práci.

Prvními aplikovanými metodami jsou geografický střed, směrodatná vzdálenost a směrodatná elipsa. Tyto metody popisují uspořádání vybraných prvků v prostoru. **Geografický střed** je průměr souřadnic místa bydliště pacientů. **Směrodatná vzdálenost** je míra rozptylu hodnot kolem geografického středu (Netrdová [online]). **Směrodatná elipsa** odchylek vykazuje, zda prostorové rozmístění jevů vykazuje rysy směrovosti (Netrdová [online]). U jednotlivých metod byla využita možnost volby váhového parametru, tedy parametru, na který je kladen důraz. Ve všech případech se jednalo o počet překročených norem.

Pro popis uspořádání objektů v geografickém prostoru byla využita funkce **Kernel Density**, která váží počet výskytů daného jevu vzhledem k ploše. Zde byl vážen počet překročených norem na 1 km<sup>2</sup>. Tato metoda byla aplikována jak souhrnně, tak samostatně pro všechny parametry funkce štítné žlázy. Jelikož je rozmístění obyvatel nepravidelné v ploše díky koncentraci ve městech a vesnicích, byla ještě využita informace o počtu obyvatel v dané obci. Tuto informaci obsahuje mapová vrstva ArcČR k roku 2001. Na základě toho byl vypočítán počet překročených norem na počet obyvatel dané obce a tím byl eliminován vliv koncentrace obyvatel a bylo možné jednotlivé obce mezi sebou porovnat a popsat výskyt jevu. Tento postup byl rovněž aplikován jak souhrnně, tak samostatně pro všechny parametry funkce štítné žlázy.

Další aplikovanou metodou prostorové analýzy je **metoda nejbližšího souseda**. Tato metoda studuje vzdálenosti mezi body, a to mezi nejbližšími sousedními body. Podstatou této metody je porovnávání pozorované průměrné vzdálenosti mezi nejbližšími sousedy a očekávané průměrné vzdálenosti. Očekávaná průměrná vzdálenost je vzdálenost mezi nejbližšími sousedy u náhodně rozmístěných bodů (Netrdová [online]). Pokud je průměrná vzdálenost mezi sousedy menší než průměr hypoteticky lokalizované distribuce, jedná se o shluk. V případě, že je průměrná vzdálenost větší než průměrná vzdálenost hypoteticky rozmístěné distribuce, jsou analyzované prvky považovány za rozptýlené (ArcGIS Pro 2017).

**Ripleyho K – funkce** je další z aplikovaných metod. Tato funkce poskytuje přehled prostorové závislosti událostí pro široký rozsah měřítek. U Ripleyho K-statistiky jsou nejprve vytvořeny buffer zóny kolem všech bodů o velikosti zvoleného prostorového kroku a následně je zjištěn počet bodů v každé zóně. Závěrečným krokem je porovnání získaného rozdělení s náhodným rozdělením četností (Netrdová, [online]). Na základě křivky hodnot je poté rozhodnuto, jestli se jedná o shlukové nebo náhodné uspořádání.

Následnou metodou prostorové analýzy bylo **Lokální Moranovo I**, které identifikuje prostorové seskupení prvků s vysokými nebo nízkými hodnotami. Součástí metody je identifikace prostorových odchylek. Pro tento účel nástroj vypočítá místní hodnotu Moranova I, z- skóre, p hodnotu a kód reprezentující typ clusteru pro každý statisticky významný rys (ArcGIS Pro 2017). Výsledkem je vznik pole značícího typ bodu, kdy se může jednat o významné seskupení vysokých hodnot (HH), nízkých hodnot (LL) a dále odlehlosti, ve kterých je vysoká hodnota obklopena nízkými hodnotami (HL) anebo naopak, nízká hodnota obklopena vysokými hodnotami (LH).

Závěrečnou prostorovou analýzou byl **variogram**. Variogram slouží k popisu chování prostorové veličiny. Znázorňuje charakter prostorově korelované variability dat, konkrétně jejich míru nepodobnosti. Nejprve je vytvořen variogram mrak, který zaznamenává míru rozdílnosti jednotlivých bodů na základě vzdálenosti. Následně je zvolen krok variogramu, který variogram mrak rozdělí do distančních tříd. V každé distanční třídě je následně vypočítána průměrná nepodobnost, která je zanesena do grafu. Spojením bodů průměrné nepodobnosti je vytvořen empirický variogram. Posledním krokem je tvorba teoretického variogramu. Křivka teoretického variogramu je modelem průměrné nepodobnosti dvojic prostorových bodů s určitou vzdáleností a je vyjádřena jednoduchým vzorcem s několika parametry (Ježek 2016).

### 3.1.2 ArcGIS online

Pro veřejnou prezentaci jak samotných dat, tak i výsledků analýz byl zvolen produkt společnosti ESRI ArcGIS online. ArcGIS online je on-line webový portál, který umožňuje používat, tvořit mapy, scény, aplikace, analytiku ale i samotná data (ArcGIS 2018). V tomto webovém nástroji je k dispozici velké množství map zachycujících celý svět v různých úhlech pohledu, dále je zde zabezpečený cloud, který byl podstatný pro samotnou tvorbu aplikace. Předností je propojenost s programem ArcGIS Desktop, kdy nedochází k problémům se sdílením dat.

Z důvodu velkého rozsahu publikovaných dat bylo potřebné zvolit vhodný formát prezentace. Jak již bylo zmíněno, ArcGIS online nabízí velké množství nových přístupů k práci s geografickými daty, ale pro účely této diplomové práce byla zvolena tvorba pomocí aplikace Story Maps. Po zvolení této aplikace je okamžitě k dispozici možnost vytvořit svůj příběh (Story Map). V ní je možné vybrat některý z předpřipravených formátů příběhu nebo si vytvořit kompletně novou šablonu. V této diplomové práci byla zvolena šablona Map Series, která nabízí možnost vytvoření záložek v horní liště aplikace. Následně se zvolí pouze název příběhu a lze začít načítat data.

V nastavení této aplikace je k dispozici možnost změnit layout a jeho vzhled a umístění panelů, dále je zde mapové nastavení, barevnost a vzhled hlavičky. V blízkosti nastavení se nachází také možnost sdílení, náhledu a nápověda. V průběhu přípravy aplikace bylo nastaveno sdílení pouze pro komunitu lidí, která byla v ArcGIS online vytvořena. Aplikace tím nebyla přístupná ostatním uživatelům.

Následně se již přešlo k tvorbě samostatných záložek. Aplikace Story Maps je navržena spíše pro kratší úderné prezentace o doporučené velikosti do 10 samostatných odrážek, záložek či stránek. Zde ale bylo třeba přehledně prezentovat velké množství mapových podkladů. Při tvorbě nových záložek se nalézají varianty prezentace v podobě mapy, obrázku, videa a webové stránky. Zde byla zvolena možnost vložení Story Map do Story Map pomocí webového odkazu. Nejedná se o standartní postup a pravděpodobně ani vývojáři nepředpokládali toto využití. Výsledný layout v sobě zahrnuje hlavní Story Map v podobě vrchních záložek, které mají přiděleny sekce v podobě puntíků druhé Story Map. Díky tomu bylo možné přehledně a na jednom místě prezentovat veškerá podkladová data a výsledky jednotlivých analýz.

Je tedy třeba vytvořit pro téměř každou záložku, kromě úvodu a závěru (informací), samostatnou Story Map. I přesto, že tímto krokem bylo získáno velké množství prostoru

pro prezentaci jednotlivých souborů dat, nebyly tyto datové soubory vždy prezentovány odděleně. Datové soubory, které by nesplňovaly vhodnou zaplněnost mapy, byly sloučeny do jedné mapy. Zde se jednalo o datové soubory týkající se biomonitoringu zátěže přírody poskytnuté společností VÚKOZ a datové soubory zachycující kvalitu pitné vody od SZÚ. Další sekce vždy znázorňují pouze jeden jev, a to geoparametry, zdravotní parametry, parametry kvality ovzduší a prostorovou analýzu. U sekce O2PLS se předkládaná data liší podle výsledků dané analýzy.

Založení aplikace bylo stejné jako u hlavní Story Map Příbramsko, pouze byl zvolen layout Map Journal s pevným textovým panelem na straně. Každá aplikace byla pojmenována podle názvu záložky, ke které byla následně v hlavní aplikaci pomocí webového odkazu přiřazena. V této verzi aplikace je stejný vrchní panel s nastavením, sdílením, náhledem a nápovědou. Na začátku Story Map automaticky provede nového uživatele nastavením úvodní stránky, kde je možnost zvolit nejprve vyplnění hlavního panelu buď mapou, obrázkem, videem, nebo webovou stránkou. Následně se již vyplňují textové pole, tedy postranní panel. Další jednotlivé odrážky jsou přidávány pomocí možnosti přidat sekci a její vytvoření probíhá stejně jako u úvodní sekce. Tyto odrážky je poté možné volně přesouvat v rámci jedné aplikace nebo dále editovat.

Postranní panel slouží primárně pro textové účely a jeho možnosti jsou poměrně rozsáhlé. Je možné nastavit styl, velikost nebo barvu textu nebo uvést text do původního stavu. Do textu je poté možné nahrát obrázky, videa a webové stránky. V této aplikaci byla několikrát využita možnost vložení obrázků s automaticky odlišeným popisem. Zároveň byla zahrnuta možnost zvětšení obrázku při kliknutí. Dalšími nabídkami editace textu v aplikaci jsou standardní odrážky a zarovnání. V aplikaci bohužel chybí možnost tvorby horních a dolních indexů, které byly potřebné pro popis znázorňovaných dat.

Pod názvem Story Actions v editaci textu se nacházejí další nástroje, ke kterým je vložena i nápověda. Jedná se o funkce, které usnadní uživateli orientovat se ve vytvořené aplikaci při odkazování na jiný text nebo na konkrétní místo. V aplikaci je několikrát odkazováno na mapy, které jsou použity již v jiné sekci, a nebylo by nejefektivnější je prezentovat dvakrát. Z tohoto důvodu zde byla využita možnost změny hlavní plochy, kde se jednoduše nastavila vybraná mapa, která se zobrazí při kliknutí na označený text. Mapu je poté možné zavřít křížkem v pravém horním rohu. Další možností, která byla v aplikaci použita pro zpřehlednění prezentace, je lokalizace adresy nebo místa, kde byl označen název obce nebo místa a pomocí tohoto nástroje byl textu přiřazen bod v mapě. Názvy obcí nebo míst jsou formátově odlišeny od ostatního textu a při kliknutí na něj se hlavní mapa přiblíží na dané místo. Posledním

aplikovaným nástrojem je možnost sekci změnit. Text je opět odlišen a při jeho zvolení je uživatel navigován do jiné sekce, která je předem nadefinována tvůrcem.

V jednotlivých aplikacích bylo pro vyplnění hlavního panelu využito obrázků nebo map. Obrázky byly použity tam, kde text pojednává o původu dat nebo následně zhodnocuje výsledky jednotlivých částí, tedy úvodní a závěrečné sekce. U obrázků je variabilita vyplnění hlavního pole tak, aby vždy bylo vidět, co je požadováno. Tvorba map tak, aby byly funkční pro uživatele aplikace, je složitější než pouhé nahrání obrázku nebo vložení odkazu na webovou stránku.

Před tvorbou samotných map je vhodné nahrát do ArcGIS online vrstvy obsahující data, která mají být vizualizována. Příprava vrstev pro nahrání je popsána v části práce věnující se programu ArcMAP. Vrstvy je možné nahrát přímo při tvorbě jednotlivých map. Poté jsou vrstvy vázány přímo na konkrétní mapu a není možné je využít opakovaně. Zároveň tento postup výrazně navyšuje velikost aplikace a přispívá k jejímu delšímu načítání, které může být pro uživatele odrazující. Druhou možností je nahrát vrstvy přímo v části s názvem obsah volbou přidat položku. Tato volba umožní přístup k daným vrstvám pro tvorbu více map a zároveň není tak objemově náročná pro prohlížení aplikace. Z těchto důvodů byla v diplomové práci zvolena druhá možnost. Při nahrávání je zvoleno, odkud se daná položka nahraje. V této diplomové práci byly vždy nahrávány vrstvy z počítače, které byly uloženy ve formátu .zip. Dále byl identifikován typ souboru Shapefile, který následně byl pojmenován. K souboru je nutné přidat tagy pro usnadnění orientace mezi jednotlivými soubory nebo pro případné vyhledávání dalšími osobami, je-li zvolena možnost uložení tohoto souboru jako veřejná. Takto byly postupně nahrány všechny vrstvy, které jsou v aplikaci prezentovány.

Tvorba samotné mapy může probíhat dvěma způsoby. Prvním způsobem je v obsahu zvolit možnost vytvořit a následně vybrat mapu. Druhý způsob probíhá přímo v aplikaci. V této diplomové práci byl použit druhý způsob. V aplikaci při editaci hlavní plochy se nachází možnost mapa, v níž se objeví okno pro zvolení dané mapy a tím možnost vytvořit mapu. Následně se vytvoří nový prostor pro tvorbu. Při tvorbě první mapy byla nastavena online podkladová mapa, která je výchozí pro všechny následně vytvořené mapy. Poté již byly přidávány jednotlivé vrstvy mapového pole díky volbě přidat a hledat vrstvy v mém obsahu. Zde byly vždy zvoleny vrstvy odpovídající danému obsahu sekce. U jednotlivých vrstev se vždy nastavuje styl zobrazení jednotlivých parametrů. V aplikaci je využita bodová metoda, kartogram, liniová metoda a mnohé další nástroje zobrazení dat. Aplikace má širokou základnu možností zobrazení jednotlivých atributů dané vrstvy a její nastavení je poměrně intuitivní a jednodušší než u programu ArcMap.

Dále byl nastaven název vrstev, které se budou zobrazovat v legendě. U samotné vrstvy byla hojně využívána možnost nastavení vyskakovacích oken. Vrstvy zde používané často obsahovaly veškeré informace, které byly k danému místu k dispozici. Pro přehlednost ale byly regulovány právě v nastavení vyskakovacích oken tak, aby se vždy zobrazila pouze informace, která se konkrétně váže k dané sekci. V tomto nastavení byly také měněny názvy jednotlivých polí. Program ArcMap, ve kterém byly jednotlivé vrstvy vytvářeny, má pouze omezené množství znaků pro pojmenování polí. Z toho důvodu bylo nutné názvy změnit tak, aby uživatel věděl, co daná hodnota představuje. Toto bylo nutné změnit u každé vrstvy. Pokud již byly ve vrstvě přejmenovány sloupce atributové tabulky, bylo možné takto upravenou vrstvu uložit. Informace o přejmenování se taktéž uložily. Pak již nebylo nutné vše měnit znovu. Tato možnost byla využita při tvorbě map v sekci Prostorová analýza. Dalšími funkcemi, které jsou zde k dispozici, je filtr a tabulka, ty ale v diplomové práci byly využívány minimálně.

Takto vytvořená mapa byla uložena a aplikace se automaticky vrátila k nastavení hlavního panelu. Zde byla nastavena pozice jednotlivých map tak, jak je má uživatel vidět při otevření aplikace. Dále je možné ještě pomocí obsahu zvolit, jaké vrstvy chce uživatel konkrétně zobrazit a jestli má být vyskakovací okno jednotlivých prvků otevřeno. Tyto nástroje však v aplikaci využity nebyly. Poslední zde využitou volbou bylo zobrazení legendy, která uživatele podrobně informuje o významu zobrazení jednotlivých vrstev na dané mapě. Po veškerém nastavení mapových podkladů bylo vše uloženo.

Takto postupně vzniklo šest aplikací: geoparametry, zdravotní parametry, parametry kvality ovzduší, parametry biomonitoringu, prostorová analýza a O2PLS. Aplikace byly uloženy ve složce můj obsah na stránkách ArcGIS online. U každé aplikace byl URL odkaz, který byl využit při tvorbě finální Story Map, která jednotlivé aplikace sdružuje pod jednu. Tato konečná Story Map byla pojmenována Příbramsko. Zde byla vždy vytvořena nová záložka a byl zde použit URL odkaz, který následně aplikace propojil. Výhodou tohoto postupu bylo, že díky propojení bylo možné jednotlivé aplikace stále upravovat a úpravy se projeví i ve finální aplikaci Příbramsko. Nebylo tedy nutné vždy nahrávat nový URL odkaz.

Posledním krokem bylo nastavení možnosti sdílení. Jelikož byla veškerá zdravotnická data anonymizována, nehrozí již jejich zneužití a je tedy možné aplikaci zpřístupnit veřejnosti. Nastavení sdílení je v hlavním obsahu při volbě dané aplikace v části nastavení sdílení, kde byla zvolena možnost veřejné.

Výsledkem tohoto postupu byl vznik aplikace, která přehledně, jednoduše, dostupně a zajímavě prezentuje data týkající se životního prostředí a medicínální data v zájmovém území okresu Příbram. Aplikace bude zveřejněna po obhajobě diplomové práce na tomto odkazu:

<https://cuni.maps.arcgis.com/apps/MapSeries/index.html?appid=82411bd3e15a4fd79fb3f24bf82d6974>.

### 3.2 Obousměrné ortogonální projekce do latentní struktury (O2PLS/OPLS)

Metoda obousměrných ortogonálních projekcí do latentní struktury (O2PLS) je variantou z metody parciálních nejmenších čtverců (PLS), či správněji projekcí do latentní struktury. Jedná se o vícerozměrnou regresi s redukcí dimenzionality, která vysvětluje vztahy mezi maticí predikovaných proměnných  $Y$  a maticí prediktorů  $X$ , a to dokonce i v obráceném pořadí (Trygg a Wold 2003).

O2PLS pomocí lineárních kombinací rozkládá matice  $X$  a  $Y$  tak, aby korelace mezi těmito maticemi byla maximální. Zároveň zajišťuje snadnou interpretovatelnost vzniklého modelu. O2PLS vytváří dvě nezávislé, ortogonální skupiny komponent, a to komponenty prediktivní sdílející rozptyl jak s prediktory v matici  $X$ , a zároveň vysvětlující strukturu vztahů mezi  $X$  a  $Y$ , tak s predikovanými proměnnými v matici  $Y$ . Na druhé straně komponenty ortogonální, vysvětlující variabilitu uvnitř matice  $X$  nezávisle na matici  $Y$  a variabilitu uvnitř matice  $Y$  nezávisle na matici  $X$  (Trygg a Wold 2003). Zatímco komponenty prediktivní jsou cílem analýzy a bývají interpretovány, komponenty ortogonální jsou z hlediska cílů O2PLS nepodstatné. Nicméně i ortogonální komponenty jsou zpravidla nezbytnou součástí celého O2PLS modelu.

V prvním kroku analýzy jsou ověřeny předpoklady normality rozdělení proměnných a reziduí, jejich konstantního rozptylu (homoskedasticity) a homogenity rozdělení dat a reziduí. Teprve po splnění těchto předpokladů je možné data analyzovat. Data, která nevyhovují výše uvedeným předpokladům, jsou obvykle normována. Od měřené hodnoty je odečten průměr a výsledek je vztažen ke směrodatné odchylce. Až takto upravená data mohou být podrobena analýze.

Metoda O2PLS, podobně jako jiné varianty PLS, je způsobilá pracovat s daty, která mohou být zatížena vysokým stupněm multikolinearity uvnitř matic  $X$  a  $Y$ . Na rozdíl od regresních metod bez redukce dimenzionality (např. vícenásobná regrese), které multikolinearita deformuje, v případě O2PLS multikolinearita zesiluje prediktivitu a při jejím dostatečně vysokém stupni si lze představit i fungující modely O2PLS o počtu proměnných převyšujícím počet měření či zahrnutí proměnných, která jsou ve vzájemném funkčním vztahu (Stramová 2017).

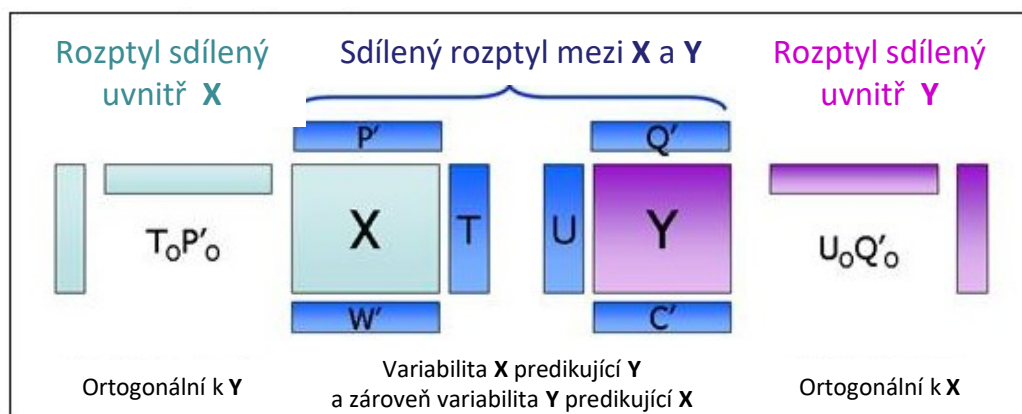
O2PLS model lze vyjádřit takto (vzorec 1 a 2):



$$X = T_p \cdot P'_p + T_0 \cdot P'_0 + E \quad (1)$$

$$Y = U_p \cdot Q'_p + U_0 \cdot Q'_0 + F \quad (2)$$

Kde  $X$  je matice s  $k$  prediktory a  $i$  subjekty,  $Y$  je matice s  $l$  závisle proměnnými a  $i$  subjekty.  $T_p$  je matice komponentních skóre prediktivních komponent extrahovaných z matice  $X$  a  $T_0$  je matice komponentních skóre komponent ortogonálních extrahovaných z matice  $X$ .  $P_p$  je matice komponentních vah prediktivních komponent extrahovaných z matice  $X$  a  $P_0$  je matice komponentních skóre komponent ortogonálních extrahovaných z matice  $X$ .  $E$  představuje matici chyb (reziduí) extrahovaných z matice  $X$ . Analogicky,  $U_p$  je matice komponentních skóre prediktivních komponent extrahovaných z matice  $Y$  a  $U_0$  je matice komponentních skóre komponent ortogonálních extrahovaných z matice  $Y$ .  $Q_p$  je matice komponentních vah prediktivních komponent extrahovaných z matice  $Y$  a  $Q_0$  je matice komponentních skóre komponent ortogonálních extrahovaných z matice  $Y$ .  $F$  představuje matici chyb (reziduí) extrahovaných z matice  $Y$  (Trygg, Holmes a Lundstedt 2007; Stramová 2017).



**Obr. 4 O2PLS modelující dvě datové tabulky**  
(Zdroj: Eriksson , Wold a Trygg [online])

Testování relevance jednotlivých prediktorů je uskutečněno statistikami Variable Importance (VIP) pro  $j$ -té proměnné  $k$ -té komponenty ( $VIP_{j,k}$ ). Metoda VIP je stanovená k odhadnutí významnosti proměnných v projekci modelu O2PLS nebo OPLS. Setrvání nebo vyloučení proměnné v modelu se určuje na základě velikosti VIP (JMP 2018). Při hodnotě VIP větší než 1 proměnná v modelu zůstává, při hodnotě VIP menší než 1 je proměnná méně podstatná a pravděpodobně je vhodným kandidátem na odebrání z modelu. Při rozhodování o zachování nebo vyřazení vždy záleží na skladbě dat a zachování patrného trendu v datech. VIP hodnota je vždy jiná po předělení modelu. U Metody O2PLS byla zvolena možnost ponechání prediktoru jako relevantního pokud je alespoň jedno z kritérií  $VIP_{j,k}$  významné,

čímž však není ještě zaručena významnost komponentních vah prediktivních komponent tohoto prediktoru. Proměnná může být relevantní pro výstavbu modelu, avšak zároveň vlivná či ne ve vztahu ke komponentním vahám prediktivních komponent. Relevance komponent byla hodnocena metodou parciálních reziduálních součtů čtverců (Partial Residual Sum of Squares, PRESS).

Model O2PLS je schopný vyrovnat se s mírným množstvím chybějících údajů, pokud jich není více jak 50 % v daném řádku či sloupci matice  $X$  či  $Y$  (Madsen, Lundstedt a Trygg 2010). Analýza O2PLS byla vypočítána s využitím software SIMCA-P+ verze 12.0.0.0 od firmy Umetrics AB (Umeå, Švédsko).

Po samotné analýze ještě probíhalo zpracování výsledků a byla vypočítána křížová validace (Cross Validation). Jedná se o metodu, která zjišťuje ovlivnění nezávislého vzorku dat výsledky statistické analýzy. Křížová validace rozdělí soubor dat do dvou skupin. Jedna skupina je trénovací a druhá skupina je určena pro testování klasifikátoru. Trénovací soubor je používán při tvorbě klasifikátoru (Malá 2013). Testovací soubor následně slouží k ohodnocení klasifikátoru. Trénovací a testovací soubory nemají mít ani jeden společný prvek. Data sloužící k hodnocení klasifikátoru tedy nejsou přítomna při tvorbě klasifikátoru. Tento postup je pak při analýze několikrát opakován. Vyhodnocení klasifikace je bráno jako průměr všech chyb predikce na rozdílných testovacích souborech (Clarke a kol. 2009).

## 4 Případová studie – Příbramsko

Tato kapitola podrobně popisuje případovou studii Příbramsko. V úvodu kapitoly je představeno zájmové území a důvod k jeho výběru. Následně jsou podrobně popsána a graficky prezentovaná veškerá data, která vstupovala do jednotlivých analýz. Poté následuje popis aplikovaných prostorových analýz. V závěru kapitoly jsou představeny výsledky analýzy O2PLS. Veškeré texty, které tato kapitola obsahuje, byly použity při tvorbě webové aplikace pro veřejnost, která bude k dispozici po obhájení práce na tomto odkazu:

**<https://cuni.maps.arcgis.com/apps/MapSeries/index.html?appid=82411bd3e15a4fd79fb3f24bf82d6974>**.

V hlavní textové části nejsou uvedeny veškeré použité texty, ale pouze ukázkové příklady. Zbylé texty, které byly použity v aplikaci, jsou v příloze 1.

### 4.1 Popis zájmového území

Zájmové území Příbramsko bylo vybráno z důvodu nezbytných požadavků na data a jejich dostupnost. Data musela zahrnovat zájmové území, charakterizovat funkce štítné žlázy, charakterizovat prostředí okresu Příbram, ve stejné nebo blízké době jako byla získána data o funkci štítné žlázy a být volně k dispozici z internetu nebo po dohodě s organizací pořizující data. Data charakterizující životní prostředí, jsou k dispozici pro celou Českou republiku. Jsou k dispozici na internetu nebo knižně volně s podmínkou citace od Českého hydrometeorologického ústavu, Výzkumného ústavu Silva Taroucy pro krajinu a okrasné zahradnictví, Ředitelství silnic a dálnic, Státního zdravotního ústavu a České geologické služby. Omezením byla zdravotnická data. První sběr zdravotnických dat, který zkoumají riziko vzniku thyreopatií, probíhal od roku 1991 do roku 2002 po celé České republice. Druhá zdravotnická vyšetření byla realizována v roce 2005, a to pouze na třech místech republiky. Jednou z těchto oblastí bylo právě Příbramsko, které bylo zvoleno díky svému jedinečnému geologickému podkladu a průmyslových závodů celorepublikové významnosti. Tato data zároveň splňovala podmínku dostupnosti, která byla vyjednána s vedením Endokrinologického ústavu.

Okres Příbram se nachází v jihozápadní části Středočeského kraje. Hlavním sídlem je město Příbram. Rozloha okresu je 1 563 km<sup>2</sup> a z toho 47,6 % plochy zabírá zemědělská půda. V okrese žije 114,2 tisíc obyvatel. Hustota zalidnění je 73,1 obyvatel na km<sup>2</sup>. Tím se řadí mezi tři

nejřidčeji osídlené okresy v kraji. Okres je dělen na tři správní obvody (Příbram, Sedlčany a Dobříš) a jeho součástí je 120 obcí (ČSÚ 2017).

Příbramsko je členitým regionem s významnou částí zalesněné krajiny, která zabírá plochu 40,4 %. Na západě a severozápadě se rozprostírá Chráněná krajinná oblast Brdy, která tvoří přirozenou hranici regionu. Ze širokých údolí krajina pozvolně stoupá v táhlé plošiny a hřebeny. Nachází se zde nejvyšší hora středních Čech Tok (865 m n. m.). Významný je také vrch Praha (862 m n. m.) a Třemšín (827 m n. m.). Nejnížší bod je v blízkosti soutoku říčky Křeničná s Vltavou s výškou 271 m n.m.. Oblastí protéká řeka Vltava, která jí dělí na dvě nestejně velké části. Nachází se zde přehrada Orlík, která je co do objemu vody největší přehradou v České republice, a přehrada Kamýk (ČSÚ 2017, Stramová 2017).

Příbramsko je díky dálnici D4 napojeno na hlavní město Prahu. Územím dále prochází silnice I. třídy - I/4, I/18, I/19 a I/66. Vede zde také důležitá železniční trať 200 Zdice – Protivín. Další železniční tratě mají spíš lokální význam (Stramová 2017).

Příbramsko je oblíbenou turistickou destinací. Nalézají se zde dvě výjimečná území, a to chráněná krajinná oblast Brdy a národní přírodní rezervace Drbákov - Albertovy skály. Jedná se o jeden z nejzachovalejších ekosystémů pro střední Vltavu, ve kterém jsou zajímavé skalní srázy a zachovalé suťové lesy s hojným výskytem tisu. Dále je na Příbramsku celá řada drobných přírodních a významných kulturních památek. Z těchto důvodů je Příbramsko oblíbeným turistickým cílem.

Velmi častým návštěvnickým cílem na Příbramsku je Hornické muzeum Příbram, které vzniklo díky těžbě, která zde probíhala. Historie území je velmi úzce spjata s hornickou činností, o které jsou první zmínky již ve 14. století. Na počátku těžby se zde těžilo hlavně stříbro, které postupně vystřídalo železo a olovo. V 19. století zde probíhala intenzivní těžba uranové rudy (Ježek 2007). Ta přinesla do současnosti ekologické problémy, podobě pozůstatků po této těžbě. Na území okresu vzniklo i mnoho hald, tak i celé oblasti, které jsou poddolované. Tyto haldy jsou vidět na Obr. 5 Podzemní prostory jsou zatopené vodou a jsou u nich sledované normy překročení jednotlivých kontaminantů. Tyto vody jsou čerpány a procházejí čistíčkami odpadních vod, aby bylo zabraňováno jejich nekontrolovanému odtékání. Pročištěná voda se poté pouští do Dubeneckého potoka a Kocáby. Hald je na území okresu 25 a představují 27 m<sup>3</sup> hlušiny, která je kontaminována. Starší haldy jsou kontaminovány více než mladší, kdy již byla lepší technika separace uranové rudy. Většina odvalů je pouze částečně zakryta vegetací. Proto při deštích voda haldou protéká k základům, odkud pokračuje dále do podpovrchové vody nebo povrchových v okolí hald (Hodrmant 2011,

DIAMO 2018). Tyto pozůstatky těžby s sebou nesou výraznou ekologickou zátěž na své okolí, proto zde působí podnik DIAMO.

DIAMO je státní podnik, který realizuje zahlazování následků hornické činnosti po těžbě uranu, rud a uhelného hornictví v České republice. Na území okresu Příbram působí odštěpný závod Správa uranových ložisek, který zajišťuje správu likvidovaných těžebních a úpravárenských kapacit uranových a ostatních rudních ložisek na území Čech. Tato společnost provozuje čističky podzemních vod ze zatopených oblastí po uranové těžbě. Dále se zabývá zpracováním vybraných odvalů na drcené kamenivo a kal, který je následně ukládán do odkladiště (DIAMO 2018). V současné době je otázka hald v zájmovém území velmi aktuální a probíhají debaty s odborníky i občany města o řešení tohoto problému.



**Obr. 5 Haldy na Příbramsku**  
(Zdroj: Jiří Jiroušek 2018)

V Příbrami je velice rozvinutý strojírenský a potravinářský průmysl. V dané oblasti působí společnost Kovohutě nástupnická a.s., jež představuje hutnický závod, který recykluje odpady s obsahem olova, a to zejména olověné baterie. Dalším zájmem firmy je výroba olova a jeho slitin a z nich různé výrobky. V neposlední řadě je zpracovatelem odpadů s obsahem drahých kovů a odpadů elektrických a elektronických zařízení vyřazených z provozu. Tento podnik vlastní certifikát s názvem „Odborný podnik pro nakládání s odpady“, který je schválen parlamentem ČR. V roce 2002 se společnost připojila k Mezinárodní deklaraci o čistší výrobě UNEP. Tento podnik dlouhodobě spolupracuje na projektu Školy, kde se jedná o spolupráci v oblasti ekologie a odpadů (Kovohutě Příbram 2018). I přes veškerá opatření je koncentrace olova v ovzduší a půdách vyšší než je běžné. Tyto koncentrace jsou ale v rozmezí, které je

povoleno. V půdách jsou koncentrace vyšší v důsledku historie podniku. Jejím nakládání s odpadem a emisemi, které se od roku 1993 výrazně zlepšilo díky zavedení nových technologií (Jaňour a kol. 2001).

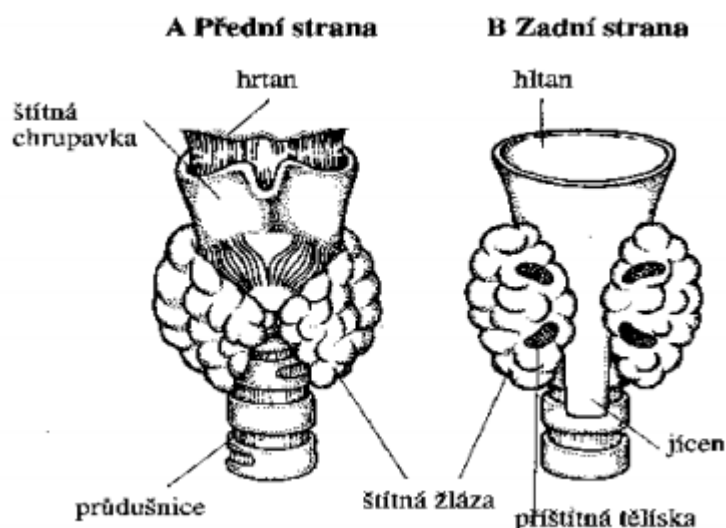
## 4.2 Vstupní data

### 4.2.1 Endokrinologie

Endokrinologie je část medicíny zabývající se diagnózou a léčbou hormonálních poruch. Tato práce se konkrétně zaměřuje na poruchy tvorby hormonů způsobené štítnou žlázou.

Endokrinní systém je kontrolním systémem endokrinních žláz, které vylučují hormony. Ty cirkulují v těle pomocí krevního oběhu a ovlivňují vzdálené orgány. Účinky vylučovaných hormonů je nezbytné kontrolovat. Pokud by byl účinek moc dlouhý nebo příliš intenzivní, mohl by způsobit narušení metabolické rovnováhy. Řídící mechanismus pracuje na principu zpětné vazby, kdy je sledován výsledný efekt procesu. Na základě zjištěných údajů se upravují faktory, které akci iniciovaly (Lidské tělo, 1996).

Štítná žláza je součástí endokrinního systému. Jedná se o důležitý lidský orgán, který produkuje hormony ovlivňující rychlost látkové přeměny. Hormony zvyšují spotřebu kyslíku a mají vliv na růst a vývoj jedince. Štítná žláza udržuje optimální úroveň metabolismu v tkáních lidského těla pro jejich normální funkci (Baranovičová 2015). Vzhled štítné žlázy je patrný z Obr. 6



**Obr. 6 Štítná žláza**  
(Zdroj: Baranovičová 2015)

Funkce štítné žlázy je řízena hormonem thyreotropin (TSH). Ten se tvoří v předním laloku hypofýzy. Slovo hormon je odvozené od řeckého slova „hormao“, což znamená „řídím“.

Hormony jsou chemické struktury, které se vyznačují vysokou účinností i v malých koncentracích. Základní jednotkou je folikul. Zde dochází k vychytání jódu, uskladnění a tvorbě dvou hormonů - thyroxinu (T4) a trijódthyroninu (T3). Jejich uvolňování stimuluje Thyreotropní hormon (TSH). Bez těchto dvou hormonů nemůžeme dlouhodobě přežít. Štítná žláza produkuje ještě třetí hormon kalcitonin, který udržuje hodnotu vápníku v těle. Aby štítná žláza produkovala hormony, potřebuje dostatečné množství jódu, který se vyskytuje v potravě. Doporučená denní dávka jódu je přibližně 150 - 200 µg/l (WHO 2001). Dalším ukazatelem funkce štítné žlázy je její objem.

## **4.2.2 Endokrinologická data**

V okrese Příbram probíhalo endokrinologické vyšetření v letech 2000 a 2005. Osoby, u kterých šetření proběhlo, byly vybrány náhodným výběrem z registru obyvatel jednotlivých regionů. Kritériem pro výběr bylo bydliště v regionu a věková kategorie. Dětská věková kategorie byla 6 - 17 let. Bylo vybráno 270 chlapců a 270 dívek. K vyšetření se dostavilo v roce 2000 291 osob a v roce 2005 274 osob.

Dodaná data byla roztríděna podle délky pobytu v daném místě a pacienti s pobytem kratším než pět let byli z výzkumu odebráni. Následně byli odebráni pacienti, u kterých nebylo možné zjistit adresu, v neposlední řadě byly odstraněny duplicity v záznamech na základě ID pacienta. Celkově se analyzovalo 490 pacientů. Na předkládaných mapách nejsou pacienti vyneseni přesně na místě bydliště i přesto, že adresa pacienta byla k dispozici. Jelikož se jedná o citlivá zdravotnická data, byla provedena anonymizace, která brání zneužití těchto dat.

### **4.2.2.1 Hladina jódu v těle**

V roce 1990 trpělo téměř 160 milionů osob nedostatečným zásobením jódem. Nedostatek jódu je považován za příčinu fyzické a mentální retardace. Z tohoto důvodu se v České republice začal přidávat jód do soli, zavedlo se užívání jódové soli v potravinářských výrobcích, obohacovaly se krmné směsi hovězího dobytka a drůbeže, dále se zvýšila konzumace minerálních vod a stolních vod s jódem. Toto zvyšování jódu realizovala Mezirezortní komise pro šetření jódového deficitu. Jelikož se ale jednalo o téměř neregulovatelné zvýšení příjmu jódu, nastalo riziko vzniku thyreopatií z nadměrného přísunu jódu (Čeřovská 2007).

Na mapě jsou znázorněni a barevně roztríděni pacienti podle parametru hladiny jódu v těle. Dále je možnost zvolit konkrétního pacienta a zjistit přesnou hladinu jódu v těle.

V Tab. 1 jsou zaznamenány počty překročených norem hladiny jódu v těle.

	Hladina jodurie v těle			
	Normální	Snížená	Zvýšená	Neměřeno
Počet pacientů	331	69	78	12

**Tab. 1. Hladina jodurie v těle**

#### **4.2.2.2 Objem štítné žlázy**

Objem žlázy může být zvětšen, což může způsobit útlak životně důležitých útvarů v oblasti krku, nebo se objevuje tvorba uzlů, za kterými se může skrývat i nádorové bujení. Normy objemu štítné žlázy se postupně mění s věkem, kdy žlázy přirozeně roste. V Tab. 2 jsou zapsány normy pro daný věk.

Věk	Referenční hodnoty
6	3.5
7	4
8	4.5
9	5
10	6
11	7
12	8
13	9
14	10.5
15	12
16	14
17	16

**Tab. 2. Referenční hodnoty objemu štítné žlázy**

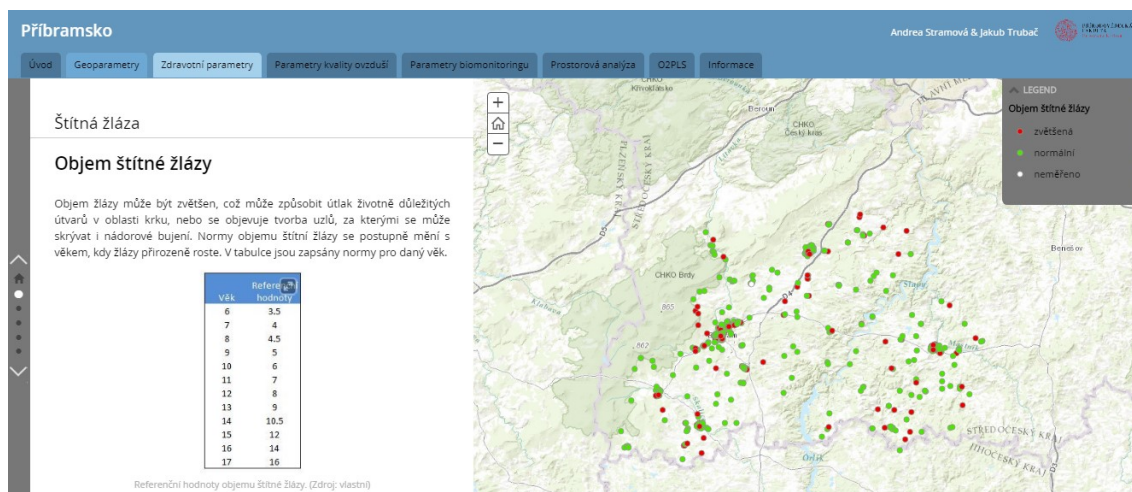
Následující tabulka (Tab. 3) poskytuje přehled počtu překročení referenčních hodnot u vyšetřených pacientů.

	Objem štítné žlázy		
	Normální	Zvětšená	Neměřeno
Počet pacientů	383	103	4

**Tab. 3. Objem štítné žlázy**

Na Obr. 7 je snímek prezentace předkládaných informací pomocí webové aplikace.





**Obr. 7** *Prezentace parametru Volgut pomocí webové aplikace*

#### 4.2.2.3 Hladina TSH v těle

Základním laboratorním vyšetřením při podezření na poruchu funkce štítné žlázy je stanovení koncentrace tyreoidálního stimulačního hormonu - TSH v krvi. TSH je hormon předního laloku podvěsku mozkového (adenohypofýzy). Jeho koncentrace se zvyšuje tehdy, když přestává štítná žláza tvořit hormony v dostatečném množství. Je-li TSH normální, můžeme poruchu funkce štítné žlázy až na vzácné výjimky vyloučit a další vyšetření nejsou potřeba. TSH zvýšené nad 3,8-4,1 mIU/l svědčí v naprosté většině případů pro subklinickou či plně rozvinutou hypotyreózu (sníženou funkci štítné žlázy).

V tabulce (Tab. 4) jsou zachyceny počty snížených, zvýšených a normálních hladin TSH v těle pacientů.

	Hladina TSH v těle			
	Normální	Snížená	Zvýšená	Neměřeno
Počet pacientů	411	2	59	18

**Tab. 4.** *Hladina TSH v těle*

#### 4.2.2.4 Hladina FT3 v těle

Volný trijódtyroninu (FT3) je využíván při diagnostice některých poruch štítné žlázy.

Na mapě jsou barevně rozlišeni pacienti podle hodnot FT3 v těle na základě normálních hodnot (2,8 - 7,1 pmol/l). Při výběru konkrétního pacienta se zobrazí naměřená hodnota FT3.

V Tab. 5 je předložen přehled počtu překročených hladin normy u FT3.

	Hladina fT3 v těle			
	Normální	Snížená	Zvýšená	Neměřeno
Počet pacientů	382	5	83	20

**Tab. 5. Hladina FT3 v těle**

#### **4.2.2.5 Hladina FT4 v těle**

V některých případech diagnostiky je vhodné doplnit i vyšetření sérové koncentrace vlastního hormonu štítné žlázy (volného tyroxinu FT4). Vzácně je vyšetřován i volný trijodtyronin (FT3).

Na mapě jsou barevně rozlišeni pacienti podle hodnot FT4 v těle na základě normálních hodnot pro dětskou populaci (12 - 24 pmol/l). Při výběru konkrétního pacienta se zobrazí naměřená hodnota FT4 a ID pacienta.

V Tab. 6 je zaznamenán přehled počtu pacientů s překročenou hladinou normy u FT4.

	Hladina fT4 v těle			
	Normální	Snížená	Zvýšená	Neměřeno
Počet pacientů	460	10	2	18

**Tab. 6. Hladina FT4 v těle**

### **4.2.3 Geoparametry**

Pod pojmem geoparametry nebo geofaktory jsou rozuměny geologické faktory, které jsou součástí životního prostředí stejně jako voda nebo ovzduší. Konkrétněji to znamená horninové prostředí s jeho komponentami a vlastnostmi významnými pro život a rozvoj společnosti včetně využitelných nerostných zdrojů (Blažková [online]). „Horninové prostředí“ je vlivným ekosystémem. Je podstatným základem a východiskem pro rozvoj organismů včetně člověka. Je rovněž nositelem nerostných zdrojů a podzemních vod.

Zdraví lidí je ovlivňováno podložím nepřímo přes fyzikální vlastnosti jednotlivých petrografických typů hornin. Při styku s radioaktivním zářením několika hornin a minerálů člověk necítí jeho zhoubný účinek bezprostředně. Jedná se o podobný efekt, jako když člověk necítí okamžitě ani účinky anomálních záření jiných fyzikálních polí - např. magnetického, elektrického, gravitačního, seismického a dalších (Blažková 1996). O škodlivých účincích radioaktivního záření nikdo nepochybuje a nebezpečí plynoucí z ozáření jsou respektována i zákonem č. 18/97 Sb. a vyhláškou č. 184/97 Sb.

#### **4.2.3.1 Geologické podloží**

Na území okresu Příbram vystupují dva základní typy horninového prostředí, sedimentární horniny Barrandienu v západní části zájmové oblasti a horniny středočeského plutonického komplexu východně od Příbrami. V prostoru dnešních Brd a Příbrami se v období spodního kambria rozkládalo rozsáhlé sladkovodní jezero, na jehož dně se usazovaly písky a další hrubozrnné usazeniny, které umožnily vznik dnešním drobům, pískovcům a lepencům. Jižní, jihovýchodní a východní část oblasti tvoří hlavně magmatické horniny středočeského plutonu, který se vyskytuje na hranici Tepelsko-Barrandienské a Moldanubické jednotky. Geologickou zvláštností je výskyt rudných žil s rozsáhlými akumulacemi rud Ag, Pb, Zn, Sb, U a Fe, které podnítily rozsáhlou těžbu. Uranové a rudné doly v této oblasti patří mezi mineralogicky světově proslulé lokality. Příbramské rudné žíly poskytly více než 300 různých druhů minerálů často v kvalitních vzorcích světové úrovně (Chlupáč 2011, Kachlík 2003, Stramová 2017).

#### **4.2.3.2 Zlomy**

Oblast Příbramska je ovlivněna značnou intenzitou poruchových pásem, zvláště dvou hlubinných zlomů - klatovského a jáchymovského. Na mapě je patrné, že převážný směr zlomových struktur je SZ-JV směru. Dokumentování zlomových struktur při základním geologickém mapování je v dané oblasti důležité, jelikož představují přírodní dráhy pro emanaci radonu.

#### **4.2.3.3 Radiometrie**

Mapa zde znázorňuje hodnoty gama záření. Jedná se o celkový příkon gama záření tří prvků, a to uranu, draslíku a thoria. Geologické podloží ČR je více než ze 60 % tvořeno metamorfovanými a magmatickými horninami, které mají vyšší koncentraci uranu a tedy následně i radonu. Proto je důležité věnovat pozornost uranu v podloží, jelikož v podobě radonu pak proniká do objektů. Zvýšená koncentrace radonu v geologickém podloží může následně ovlivnit i koncentrace radonu ve stavebních materiálech přírodního původu a ve vodě, dodávané do objektů z podzemních zdrojů. Zdroje podzemní vody pro hromadné zásobování obyvatelstva jsou systematicky monitorovány na přítomnost radonu a někdy je přistupováno k tzv. odradonování (ČGS [online]). Z toho důvodu radon v podloží nejvíc ovlivňuje výslednou koncentraci radonu v objektech.

Radon (Rn-222) vzniká radioaktivní přeměnou uranu (U-238) v horninovém podloží. V jednotlivých typech horninového podloží se mění i koncentrace uranu. V horninách přeměněných nebo metamorfovaných tlakem a teplotou během dlouhé geologické doby je větší koncentrace uranu než v horninách usazených a sedimentárních, kde se setkáváme spíše s nižšími koncentracemi. Nejvyšší koncentrace uranu se obvykle vyskytují ve vyvřelých,

magmatických horninách (ČGS [online]). Vhodným příkladem jsou žuly. Radon se může šířit v horninovém podloží a následně se dostávat až do objektu díky špatné izolaci. Nejedná se ale o stabilní izotop, a proto se může dále přeměňovat na dceřiné produkty kovové povahy, které se vážou na aerosoly a dostávají se vdechnutím do lidského organismu. Radon má tedy snahu unikat z prostředí s vyšší hustotou (jako jsou horniny a zeminy), do prostředí s nižší hustotou což je ovzduší (Mucha [online]).

Vliv radonu na lidský organismus byl nejprve pozorován u horníků, kteří těžili stříbro. Součástí stříbrných žil byla uranová mineralizace a při těžbě žil s vyšším podílem uranových minerálů klesaly obsahy stříbra. Díky nedostatečné ventilaci dolů byli dělníci vystaveni vysokým expozicím a poté trpěli rakovinou plic a dýchacích cest. Pravá příčina onemocnění byla zjištěna až později a byla následně epidemiologicky prokázána i na běžné populaci, která žije v oblastech s vyšší koncentrací radonu (SURO 2018).

Na území okresu Příbram dříve probíhala těžba uranu, jelikož se ho zde nacházelo velké množství a stále je zde značné množství přítomno. Na předkládané mapě je viditelný celkový příkon gama záření (nGy/h). Severozápadní část okresu má spíše nižší příkon záření, a to do 40 nGy/h. Naopak v jihovýchodní části okresu v oblasti města Sedlčany je viditelný veliký příkon gama záření a to až 123 nGy/h a více. Tyto hodnoty záření jsou již klasifikovány jako vysoké radonové riziko. Tento příkon je také pozorovatelný v oblasti města Příbram. Na tomto území právě v minulosti probíhala těžba uranu.

#### **4.2.3.4 Radon v objektech**

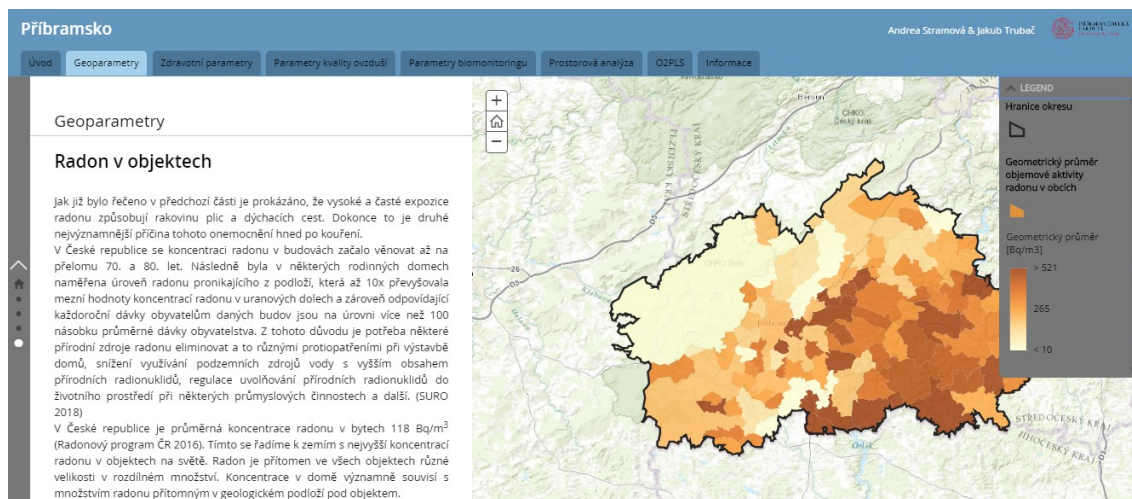
Jak již bylo řečeno v předchozí části, je prokázáno, že vysoké a časté expozice radonu způsobují rakovinu plic a dýchacích cest. Dokonce to je druhá nejvýznamnější příčina tohoto onemocnění hned po kouření.

V České republice se koncentraci radonu v budovách začalo věnovat až na přelomu 70. a 80. let. Následně byla v některých rodinných domech naměřena úroveň radonu pronikajícího z podloží, která až 10x převyšovala mezní hodnoty koncentrací radonu v uranových dolech a zároveň odpovídající každoroční dávky obyvatelům daných budov jsou na úrovni více než 100 násobku průměrné dávky obyvatelstva. Z tohoto důvodu je potřeba některé přírodní zdroje radonu eliminovat, a to různými protipatřeními při výstavbě domů, snížením využívání podzemních zdrojů vody s vyšším obsahem přírodních radionuklidů, regulací uvolňování přírodních radionuklidů do životního prostředí při některých průmyslových činnostech a další (SURO 2018).

V České republice je průměrná koncentrace radonu v bytech 118 Bq/m<sup>3</sup> (Radonový program ČR 2016). Tímto se řadíme k zemím s nejvyšší koncentrací radonu v objektech na světě. Radon je přítomen ve všech objektech různé velikosti v rozdílném množství. Koncentrace v domě významně souvisí s množstvím radonu přítomným v geologickém podloží pod objektem.

Mapa vyobrazuje geometrický průměr objemové aktivity radonu v jednotlivých obcích na základě dlouhodobého měření od roku 199 do roku 2010. Zde je vidět podobnost s předchozí mapou znázorňující hodnoty gama záření. V severovýchodní části zájmového území jsou naměřeny nízké hodnoty objemové aktivity radonu a v jihovýchodní části jsou naměřeny vysoké hodnoty. Nejvyšší průměrná hodnota byla zaznamenána v obci Kozárovice (797,6 Bq/m<sup>3</sup>) ležící právě v jihovýchodní části okresu. Při zvolení konkrétní obce naleznete informace o počtu změřených bytů, tedy počtu bytů, u kterých bylo v daném období provedeno měření. U malých obcí se stává, že zde nebyl měřen ani jeden objekt. Následně je zde informace a podílu objektů, které měli překročenou hranici 300 Bq/m<sup>3</sup> a 1 000 Bq/m<sup>3</sup>. Poslední informací zde zapsanou je přesný geometrický průměr, který je v mapě vyjádřen barvou.

Vzhled celé strany věnované tématu radonu v objektech je zachycen na Obr. 8.



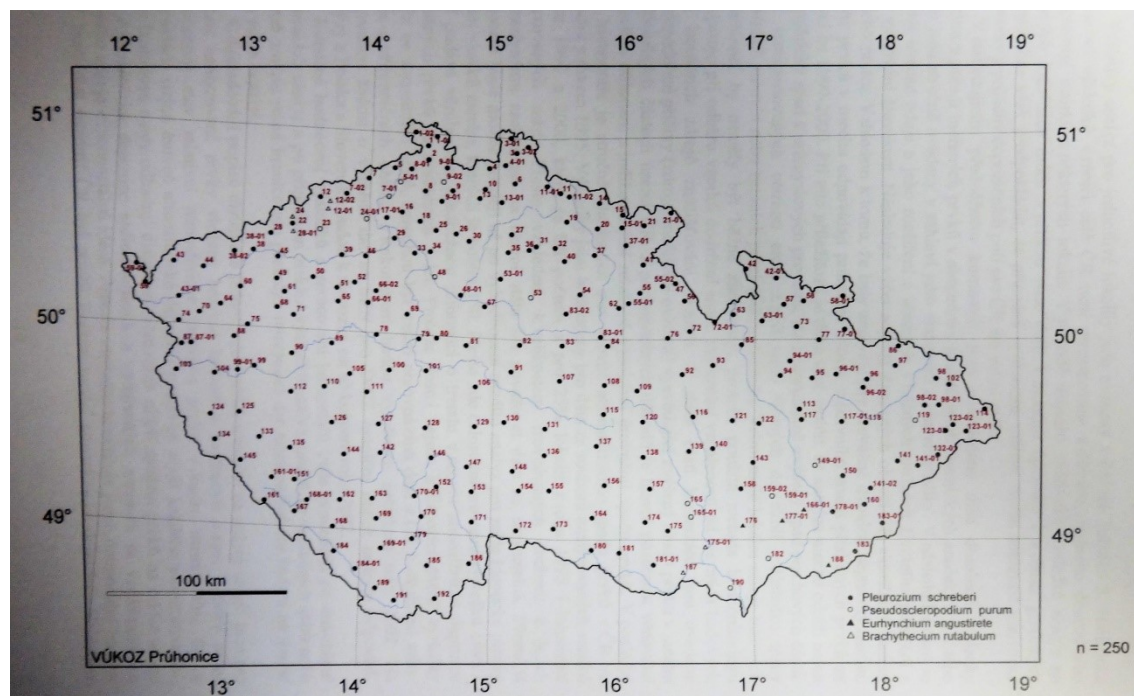
**Obr. 8** *Prezentace radonu v objektech pomocí webové aplikace*

#### 4.2.4 Data z biomonitoringu zátěže přírody a z analýz pitné vody

Bodová data prezentovaná v mapách znázorňující koncentraci jednotlivých parametrů v mechu jsou převzata od Výzkumného ústavu Silva Taroucy pro krajinu a okrasné zahradnictví (VÚKOZ). Tato veřejná instituce zkoumá všechny typy krajiny a s ní související environmentální rizika, její ochrana a další. Jedním z projektů, kterému se ústav věnuje, je Určení zátěže krajiny



ČR aktuálními úrovněmi atmosférické depozice a zajištění dat pro celoevropský biomonitorovací program 2000, (VÚKOZ 2016). Tento projekt biomonitoroval mechy a koncentraci 36 prvků v nich obsažených v roce 2000. Data pokrývají území ČR ve velmi řídkém, ale pravidelném rozmístění, které je zachyceno na Obr. 9.



**Obr. 9 Lokace sebraných vzorků mechu**  
(Zdroj: Sucharová a kol. 2004)

V okolí okresu Příbram bylo měřeno 14 vzorků a v samotném okresu byly sebrány vzorky na 5 místech. Analýza vzorků byla ve výzkumu zpracována následujícím způsobem. Na každém místě bylo celkově odebráno 9 litrů vzorku mechu. Vzorky byly následně sušeny na vzduchu, poté namlety v mlýnku s titanovým rotorem a přesáty na sítích, s velikostí ok 0,2 mm. Dále byly vzorky rozloženy pomocí mikrovlnné trouby pracující se zvýšeným tlakem (MARS 5) nejprve v kyselině dusičné ( $\text{HNO}_3$ ) a následně pak v peroxidu vodíku ( $\text{H}_2\text{O}_2$ ). Analýzování vzorků probíhalo na přístroji ICP-MS (PE, Elan 6000). Ve vzorcích mechu bylo zjištěno těchto 35 prvků: Ag, Al, As, Ba, Be, Bi, Cd, Ce, Co, Cr, Cs, Cu, Fe, Ga, In, La, Li, Mn, Mo, Ni, Pb, Pr, Rb, S, Sb, Sc, Se, Sr, Th, Ti, U, V, Y a Zn.

Pro okres Příbram byly vybrány prvky, které překračovaly normy na daném území, nebo vykazovaly abnormality. Jedná se o: Ag, Bi, Cd, Cr, Cu, Fe, In, Mo, Pb, Sb, Sn a Zn. Naměřená množství jednotlivých prvků byla uvedena v jednotkách  $\mu\text{g/g}$  (ppm). Získaná hodnota sice neodpovídá přesné koncentraci prvku v atmosféře, ale jednotlivé hodnoty jsou

porovnatelné. Jedná se o tzv. proxy data. V místech, kde je vyšší koncentrace prvku v mechu, je vysoce pravděpodobné, že je i vyšší koncentrace v atmosféře a naopak.

Poslední vrstvou, která je v mapách zaznamenána, je vrstva rozborů vzorků pitné vody. Pro hodnocení jakosti pitné vody je závazným podkladem vyhláška Ministerstva zdravotnictví České republiky č. 252/2004 Sb., která je plně v souladu s Evropskou směrnicí o jakosti pitné vody určené pro lidskou spotřebu. Pro účely této práce byly poskytnuty výsledky z rozborů od Státního zdravotního ústavu (SZU) pro 40 obcí, kde bydlí téměř 80 % vyšetřených pacientů. Státní zdravotní ústav je příspěvkovou organizací ministerstva zdravotnictví. Ústav je zřízen za účelem přípravy podkladů pro národní zdravotní politiku, pro ochranu a podporu zdraví, k zajištění metodické a referenční činnosti na úseku ochrany veřejného zdraví, k monitorování a výzkumu vztahů životních podmínek a zdraví, k mezinárodní spolupráci, ke kontrole kvality poskytovaných služeb k ochraně veřejného zdraví, k postgraduální výchově v lékařských oborech ochrany a podpory zdraví a pro zdravotní výchovu obyvatelstva. Jeho úkolem je sledovat dlouhodobé trendy šíření infekčních nemocí, ohrožení nemocí z povolání, epidemiologii drogových závislostí a zjištěné informace dále předávat orgánům ochrany veřejného zdraví (Státní zdravotní ústav 2016).

Na Obr. 10 je zachycen pohled z výšky na dopadovou plochu v Brdech, kde jsou současně zachyceny malé vodní plochy a mechový porost. Tento obrázek slouží jako strana ve webové aplikaci.



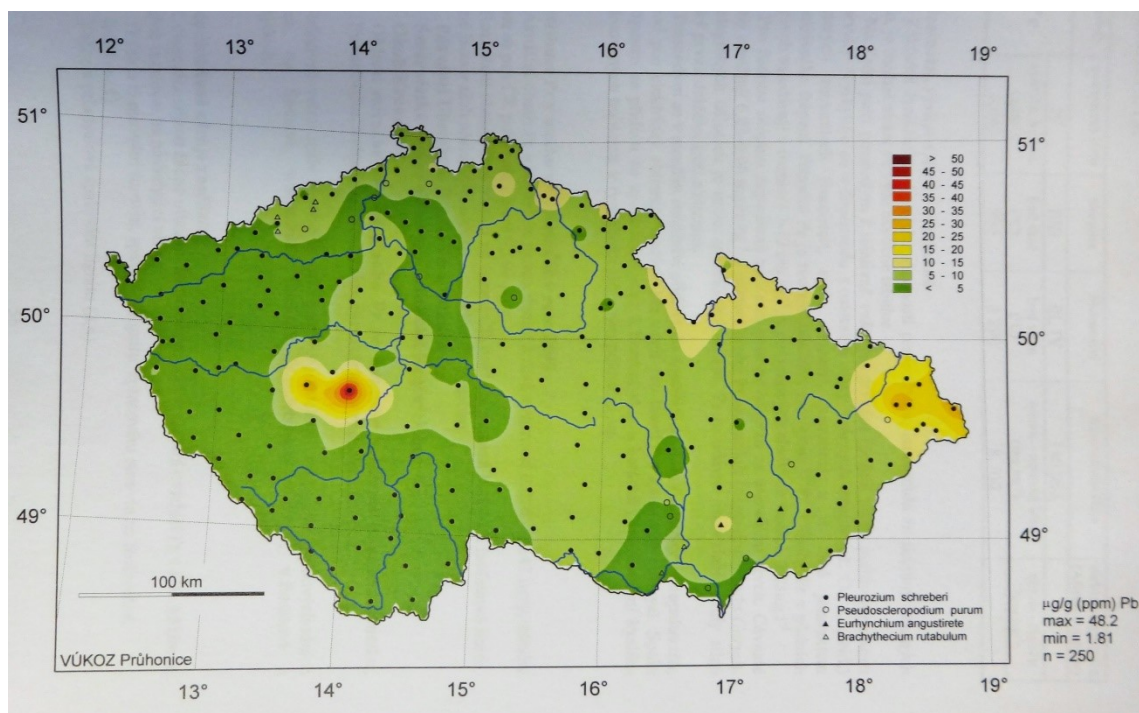
***Obr. 10 Letecký pohled na dopadovou plochu v Brdech***  
(Zdroj: Jiří Jiroušek 2018)



#### 4.2.4.1 Olovo

Na přiložené mapě jsou znázorněny koncentrace olova v mechu. Dle odhadovaného množství patří olovo mezi středně hojné prvky, které se na Zemi vyskytují. Hojnější jsou Pb-minerály než ryzí Pb. To se často vyskytuje v horninách, a to nejvíce v bazaltických vyvřelinách. Olovo není pro žádnou skupinu organismů na Zemi elementárním prvkem kromě savců, u nichž je ale význam diskutován. Všeobecně je známá toxicita Pb a jeho sloučenin, které se hromadí v různých orgánech, mohou působit anémii, poruchy centrálních nervového systému až smrt, ale to hlavně při vyšších koncentracích. Olovo má i teratogenní, alergizující a až karcinogenní účinky. Olovo má velké uplatnění v hutnictví barevných kovů, kde je často využíváno do slitin. Může být využíváno k výrobě trubek, akumulátorů, chladičů, ochraně elektrických kabelů, jako aditivum do benzínu, ale i při výrobě křišťálového skla a mnoho dalšího (Sucharová a kol. 2001).

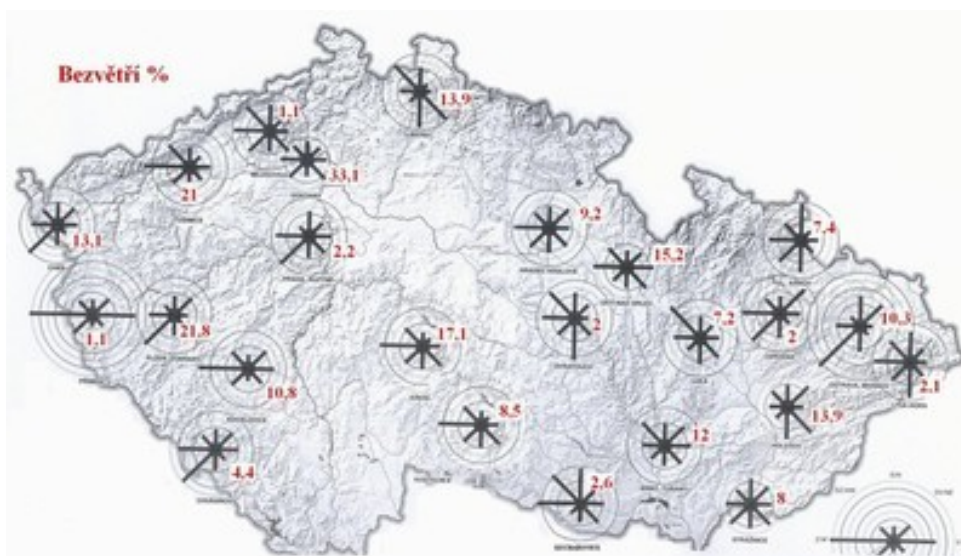
Analýza mechu pro celé území ČR zjistila obsah Pb v rozpětí od 1,81 – 48,2 µg/g. Střední hodnota obsahu Pb v mechu pro oblast je přibližně 6,75 µg/g. V celé České republice jsou identifikovány tři oblasti se zvýšenou koncentrací Pb (Obr. 11), a to Příbramsko a Rokycansko, Ostravsko a přilehlé části Moravskoslezských Beskyd, Frýdlanstko a Jizerské hory, Orlické hory, Rychlebské hory a Jeseníky. Zbylé území je pod vlivem malého obsahu Pb nebo velmi mírně zvýšeného spadu Pb.



**Obr. 11 Distribuce Pb v ČR v roce 2000**  
(Zdroj: Sucharová a kol. 2004)



Na území okresu Příbram byla největší koncentrace zjištěna v obci Trhové Dušníky (48,2  $\mu\text{g/g}$ ), která leží v těsné blízkosti okresního města Příbram. V okresním městě nebyl proveden sběr vzorků mechu. Je možné se domnívat, že zde by byla koncentrace ještě větší nebo přibližně stejná. V severní části města sídlí velké podniky, které mimo jiné zpracovávají olovo a drahé kovy, jedná se např. o Kovohutě Příbram nástupnická, a. s., Halex-Schauenberg, ZRUP, a. s. a další. Druhá největší koncentrace byla naměřena za hranicemi okresu Příbram, a to u Obce Hrádek (29,6  $\mu\text{g/g}$ ). Na území Příbramska je roční převládající směr větru východní (Obr. 12), a to je možný důvod, proč je koncentrace olova na západ od pravděpodobného zdroje. Ostatní měřené vzorky se pohybují v nízkých hodnotách do 8  $\mu\text{g/g}$ .

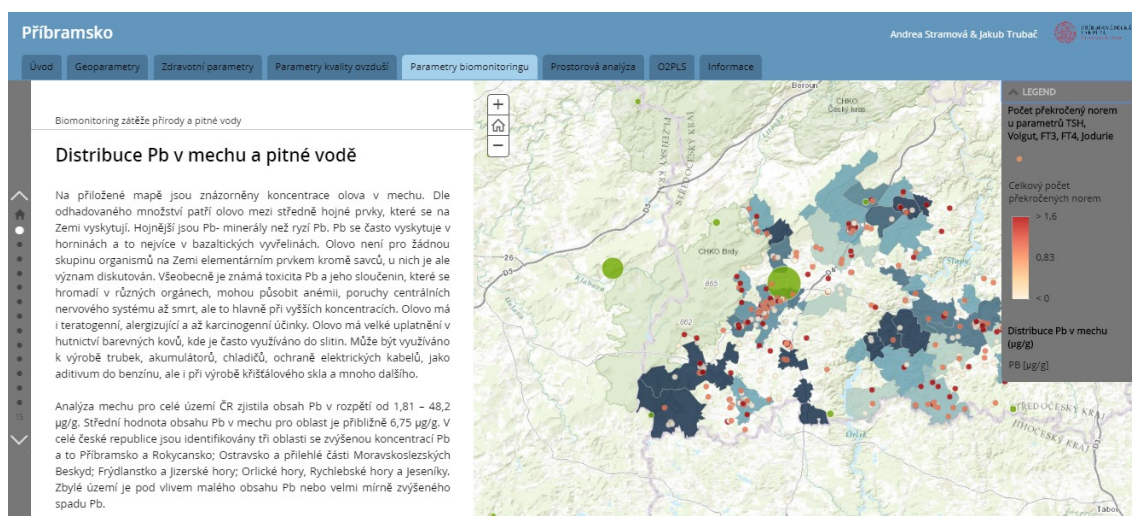


**Obr. 12 Směr větru v ČR**  
(Zdroj: ČHMÚ 2018)

Mezi těmito dvěma měřenými body, které vykazují vyšší koncentrace Pb, se rozprostírá Chráněná krajinná oblast Brdy. Nebydlí zde žádní lidé. Nejbližší je okresní město Příbram. Zde můžeme pozorovat větší množství pacientů s překročenými normami u jednotlivých vyšetření. Informaci o počtu překročených norem na jednoho vyšetřeného pacienta sděluje také bodová vrstva, která pokrývá celý okres Příbram, právě mimo CHKO Brdy z důvodu neobydlenosti této oblasti. Na druhou stranu se jedná také o město, kde žije nejvyšší počet obyvatel, a proto je možné tento vliv tady očekávat.

Nejvyšší koncentrace olova v pitné vodě (3  $\mu\text{g/l}$ ) na území okresu Příbram byla zaznamenána v obci Vrančice, která leží ve středu okresu. Přípustná mezní hodnota olova v pitné vodě je 20  $\mu\text{g/l}$ . Koncentrace olova 2  $\mu\text{g/l}$  byla naměřena u deseti obcí, které jsou rozmístěné po celém okresu. V těchto obcích se nachází malé množství pacientů, kteří mají překročenou normu dva a vícekrát. Pacienti s překročenou normou měřených parametrů se spíše nacházejí v obcích s koncentrací olova v pitné vodě okolo hodnoty 1  $\mu\text{g/l}$ .

Na Obr. 13 je snímek prezentace předkládaných informací pomocí webové aplikace.



**Obr. 13** Prezentace parametru Pb pomocí webové aplikace

#### 4.2.4.2 Zhodnocení vlivu dat z biomonitoringu zátěže přírody na funkci štítné žlázy

Veřejné instituce VÚKOZ v roce 2000 monitorovala koncentrace prvků v mechu po celé české republice. Pro tuto práci bylo zvoleno 14 bodů v okrese Příbram a jeho nejbližším okolí. U těchto bodů byly vybrány parametry, které v této oblasti vykazovaly abnormality.

Na základě koncentrací jednotlivých prvků bylo určeno pořadí pro jednotlivé obce (1-14) od nejmenší koncentrace po největší. Poté byl proveden součet těchto pořadí za každou obec a bylo vytvořeno jedno souhrnné pořadí pro všechny naměřené prvky v mechu za jednotlivá místa sběru, které je patrné v sloupečku „Celkové pořadí“ v Tab. 7.

Nejbližší sídlo	Pořadí (od nejmenší k největší koncentraci)												Celkové pořadí
	Ag	Bi	Cd	Cr	Cu	Fe	In	Mo	Sb	Sn	Zn		
Dobřív	14	14	14	14	13	14	14	14	13	13	14	14	
Jindřichovice	11	9	6	2	6	8	5	2	2	3	11	3	
Kozárovice	2	13	4	7	2	12	9	6	8	7	3	7	
Lhotka u Berouna	8	12	7	13	12	13	12	13	10	11	10	13	
Neřežín	12	11	12	5	10	1	7	12	5	4	13	9	
Obora	5	10	2	10	3	7	8	4	6	5	5	4	
Plánice	3	2	9	3	8	5	2	7	3	2	2	2	
Podmokly	10	6	8	12	14	10	10	11	4	6	9	10	
Příbram	13	7	13	8	7	2	13	3	14	14	7	11	
Starý Smolivec	7	5	3	6	5	4	4	8	7	8	8	5	
Veletín	4	8	10	9	9	9	11	5	9	9	4	8	
Voznice	6	3	5	4	4	6	3	10	11	10	6	6	
Vráž	1	1	1	1	1	3	1	1	1	1	1	1	
Vysoký Újezd	9	4	11	11	11	11	6	9	12	12	12	12	
	nejlepší výsledek					nejhorší výsledek							

**Tab. 7.** Hodnocení distribuce prvků v mechu

V pořadí za jednotlivé prvky je již patrné, že sběr mechu u obce Dobřív měl většinou největší naměřenou koncentraci jednotlivých prvků. Zároveň na druhé straně je taktéž vidět, že nejlépe, až na jednu výjimku, dopadl vzorek sebraný u obce Vráž. Ani jedna z obcí neleží v okrese Příbram. Obce, které dopadly nejlépe v celkovém pořadí (Vráž, Plánice a Jindřichovice), leží jižním směrem od okresu Příbram. Obce, které dopadli nejhůře v celkovém hodnocení (Dobřív, Lhotka u Berouna a Vysoký Újezd) taktéž neleží v okrese Příbram, ale jsou přesně na opačné světové straně od nejlepších obcí, tedy na severu od zkoumané oblasti. Ve vybrané oblasti leží 5 míst sběru a to nejbliže k obcím Kozárovce, Neřežín, Příbram, Starý Smolivec a Voznice. Obec Neřežín neleží v okrese Příbram a místo sběru bylo provedeno v bývalém Vojenském újezd Brdy, kde nikdo nežije. Místo sběru, které bylo pojmenováno podle nejbližší obce Starý Smolivec, se nachází v obci Hvoždany, tedy v okrese Příbram. V Tab. 8 je vidět hodnocení koncentrace prvků za jednotlivé obce společně s výsledky endokrinologického vyšetření pacientů, kde byly vytvořeny 3 nové sloupce s počtem vyšetřených pacientů, s počtem pacientů, kteří mají překročenou normu parametrů funkce štítné žlázy 2 a vícekrát. Aby bylo vidět celkové porovnání mezi jednotlivými obcemi, bylo vypočítáno procentuální zastoupení těchto pacientů z celkového počtu pacientů.

Nejbližší sídlo	Pořadí (od nejmenší k největší koncentraci)												Celkové pořadí	Počet vyšetřených pacientů	Počet pacientů s překročenou normou >=2	Procento pacientů s překročenou normou >=2
	Ag	Bi	Cd	Cr	Cu	Fe	In	Mo	Sb	Sn	Zn					
Kozárovce	2	13	4	7	2	12	9	6	8	7	3	7	0			
Příbram	13	7	13	8	7	2	13	3	14	14	7	11	133	22	16.54	
Starý Smolivec	7	5	3	6	5	4	4	8	7	8	8	5	7	1	14.29	
Voznice	6	3	5	4	4	6	3	10	11	10	6	6	3	1	33.33	
	nejlepší výsledek						nejhorší výsledek									

**Tab. 8. Hodnocení vlivu koncentrace prvků v mechu na funkci štítné žlázy**

V hodnocení koncentrace prvků těchto čtyř obcí dopadla nejlépe obec Starý Smolivec (Hvoždany) a nejhůře Příbram. V procentuálním zhodnocení endokrinologických vyšetření dopadla nejlépe obec Starý Smolivec (Hvoždany) a nejhůře obec Voznice. Z tabulky je patrné, že obec, kde byla nejnižší koncentrace jednotlivých prvků, je i nejnižší procentuální zastoupení pacientů s nemocí štítné žlázy. Předpoklad, že obec s nejvyšší koncentrací prvků v mechu bude mít i nejhorší výsledky u endokrinologických vyšetření se nepotvrdila a nejhorší výsledky má obec Voznice, která leží severovýchodně od města Příbram.

Z těchto výsledků vyplývá, že zvýšená distribuce jednotlivých prvků je spíše v severní části okresu Příbram a naopak nízká distribuce prvků je v jižní části. Z výsledků v druhé tabulce by mohlo vyplývat, že taktéž pacienti, kteří žijí na severu území, mají častěji problém s funkcí



štítné žlázy než pacienti žijící na jihu území, tedy, že větší distribuce vybraných prvků vede k onemocnění štítné žlázy. Pro tento závěr jsou ale data za rozborů prvků v mechu velice řídká a nejsou dostatečně prokazatelná. Při pohledu na mapové zobrazení pacientů a jejich počtu překročených norem je spíše patrné pravidelné rozložení mimo oblast Brd, kde nikdo nežije. Tento pohled ale může podpořit možné výsledky geostatistického hodnocení.

Na Obr. 14 je patrný lesní porost v Brdech, jehož součástí je právě i mech, který sloužil jako biomonitoringový prvek zátěže přírody. Tento snímek byl použit v aplikaci pro názornou představu.



**Obr. 14** *Letecký snímek na lesní porost v Brdech*  
(Zdroj: Jiří Jiroušek 2018)

#### **4.2.4.3 Zhodnocení vlivu dat parametrů pitné vody na funkci štítné žlázy**

Státní zdravotní ústav poskytl data za 40 obcí ležících v okrese Příbram za období 2004 – 2005. Tato data obsahovala údaje o koncentracích 16 parametrů pitné vody. U 3 parametrů bylo zjištěno, že byla překročena horní mez určená Světovou zdravotnickou organizací. Jednalo se o mangan, jehož hranice byla překročena v 6 obcích, dále u 2 obcí o kadmium a u 1 obce o arsen.

Na základě koncentrací jednotlivých parametrů v pitné vodě bylo určeno pro každý parametr pořadí jednotlivých obcí v podobě čísel. Toto pořadí označuje obce od nejmenších koncentrací po koncentrace největší (1-40). Následně byl proveden součet těchto pořadí

za každou obec a bylo vytvořeno jedno souhrnné pořadí pro všechny parametry pitné vody za jednotlivé obce, které je vidět v sloupečku „Celkové pořadí“ v Tab. 9.

Název obce	Pořadí (od nejmenší k největší koncentraci)																	Celkové pořadí	Počet vyšetřených pacientů v obci	Počet pacientů s překročenou normou	Procento pacientů s překročenou normou $\times 2$ (%)
	Se	As	Pl	Cd	Pb	Ag	Sb	Be	B	Al	Ch	Mn	Cu	Ni	Hg	Na					
Bohutín	31	29	1	1	23	N	N	N	N	31	N	14	14	N	1	12	13	14	4	28,57	
Březnice	25	5	29	15	10	N	5	1	1	21	31	21	11	29	1	26	29	26	5	19,23	
Čenkov	31	5	12	7	30	N	32	1	31	3	24	35	11	32	1	6	36	2	0	0,00	
Dobříš	1	26	21	17	10	N	5	1	4	16	23	33	19	28	1	17	27	39	6	15,38	
Drevníky	31	29	1	1	23	N	N	N	N	31	N	14	14	N	1	12	13	0			
Dublovice	1	5	31	14	23	N	5	1	4	40	33	20	21	35	1	32	37	7	2	28,57	
Dubno	31	29	1	1	23	N	N	N	N	31	N	14	14	N	1	12	13	1	0	0,00	
Hluboš	31	29	1	1	23	N	N	N	N	31	N	14	14	N	1	12	13	4	0	0,00	
Hvoždany	1	5	1	19	30	N	5	1	4	18	1	27	21	3	1	4	8	7	1	14,29	
Chraštice	27	5	12	7	30	1	28	1	31	3	24	30	2	3	1	26	29	0			
Jesenice	31	34	N	19	10	N	27	1	1	13	1	38	2	1	1	4	19	5	1	20,00	
Jince	1	5	1	19	10	N	5	1	4	23	1	22	40	2	1	16	11	9	1	11,11	
Kamýk nad Vltavou	1	5	12	7	30	1	32	1	4	14	24	25	8	3	1	19	21	1	0	0,00	
Klučenice	1	40	20	19	1	N	1	1	4	3	1	7	21	3	1	28	11	4	1	25,00	
Kňovice	40	33	N	39	10	N	36	N	4	3	33	22	13	33	39	29	40	2	0	0,00	
Kosova Hora	1	2	24	19	1	N	1	1	4	37	1	19	21	3	1	22	13	6	1	16,67	
Krásná Hora nad Vltavou	1	5	N	5	10	N	25	1	4	16	36	28	6	33	1	23	22	3	1	33,33	
Malá Hraštice	1	5	1	19	30	N	5	1	4	18	1	1	21	3	1	35	10	6	2	33,33	
Mokrovraty	1	34	12	6	29	6	31	1	4	25	33	31	10	30	1	30	39	2	0	0,00	
Nová Ves pod Pleší	31	5	21	19	10	N	5	1	1	2	1	40	1	36	1	37	24	2	1	50,00	
Nový Knín	1	5	1	19	10	N	5	1	4	1	1	37	21	3	39	33	18	2	0	0,00	
Obecnice	1	5	1	19	10	N	5	1	4	35	1	3	21	3	1	1	1	4	3	75,00	
Obořiště	1	5	32	19	1	N	5	1	4	26	1	7	21	3	1	7	3	2	0	0,00	
Petrovice	1	1	26	19	9	N	4	1	4	15	1	3	21	3	1	3	2	18	6	33,33	
Příbram	1	4	25	18	21	1	24	34	4	38	22	22	19	3	1	2	33	133	22	16,54	
Rosovice	1	5	32	19	1	N	5	1	4	26	1	7	21	3	1	7	3	2	0	0,00	
Rožmitál pod Třemšínem	31	27	1	19	30	N	5	1	4	24	1	3	21	3	1	40	24	16	2	12,50	
Rybníky	1	5	32	19	1	N	5	1	4	26	1	7	21	3	1	7	3	0			
Sedlčany	1	39	28	19	23	N	5	1	4	36	1	18	21	3	1	20	26	42	8	19,05	
Solenice	25	38	27	19	10	N	5	1	4	39	1	6	21	3	1	38	32	0			
Svaté Pole	1	5	32	19	1	N	5	1	4	26	1	7	21	3	1	7	3	5	2	40,00	
Svatý Jan	27	28	12	7	30	N	28	1	31	3	24	36	6	3	1	34	38	1	0	0,00	
Velká Lečice	27	5	12	7	30	1	32	1	31	12	24	29	9	3	1	24	35	1	1	100,00	
Višňová	1	5	32	19	1	N	5	1	4	26	1	7	21	3	1	7	3	2	0	0,00	
Voznice	1	5	1	15	22	N	5	1	4	3	31	26	18	30	1	21	20	3	1	33,33	
Vrančice	1	37	N	40	40	N	35		4	18	1	2	21	3	1	25	28	3	0	0,00	
Vysoký Chlumec	1	3	23	19	1	N	1	1	4	3	1	31	21	3	1	31	9	5	0	0,00	
Zalužany	27	5	12	7	30	1	28	1	31	3	24	39	2	3	1	17	29	0			
Zduchovice	31	5	12	7	30	N	25	1	31	3	24	34	2	3	1	36	34	2	0	0,00	
Županovice	1	36	30	19	10	N	5	1	4	22	1	7	21	3	1	39	23	0			
překročená mezní hodnota koncentrace v pitné vodě																	nejhorší výsledky				
																	nejlepší výsledky				

**Tab. 9. Hodnocení vlivu pitné vody na funkci štítné žlázy**  
(Zdroj: vlastní)

V tomto hodnocení dopadla nejlépe Obecnice (111) a nejhůře obec Kňovice (334). Obecnice nejhůře dopadla při hodnocení koncentrace hliníku v pitné vodě, ale ostatní parametry vždy obsahovala v nejmenším naměřeném množství. Druhou obcí, která dopadla

nejlépe, je obec Petrovice. Tato obec měla pouze o jeden bod více v součtu pořadí (112). Obec Kňovice, která byla vyhodnocena jako obec s nejhorší kvalitou pitné vody, měla dokonce jednu překročenou horní mez obsahu prvku v pitné vodě (Sb). Mokrovraty je obec s druhým nejhorším hodnocením, ale přesto získala o 50 bodů méně.

Následně je v tabulce zaznamenáno počet vyšetřených pacientů pro dané obce a počet pacientů, kteří překročili normy parametrů funkce štítné žlázy. Z těchto údajů byl pak vypočítán poslední sloupeček, který značí procento pacientů s překročenou normou 2 a více v jednotlivých obcích. U pacienta, který má překročenou normu parametru funkce štítné žlázy 2 a vícekrát se v této práci předpokládá, že trpí některou z poruch funkce štítné žlázy. V tabulce jsou označeny obce s nejmenším a největším procentem.

Obce s nejmenším procentem vyšetřených pacientů trpících poruchou funkce štítné žlázy mají často vyšetřené pouze 2 pacienty, ale je zde i obec s 5 vyšetřenými pacienty a žádný z nich netrpí poruchou funkce štítné žlázy. Tři z těchto obcí obsadily v hodnocení kvality pitné vody třetí místo, další obce obsadily pozici ve středu, ale 5 obcí obsadilo poslední místo. Obec Kňovice byla dokonce vyhodnocena obcí s nejhorší kvalitou pitné vody v dané oblasti.

Dvě obce s největším procentem pacientů trpících poruchou funkce štítné žlázy byly jedněmi z nejlépe hodnocených obcí ohledně kvality pitné vody v daném regionu.

Z těchto výsledků je možné se domnívat, že horší kvalita pitné vody, ale přesto přípustná a většinou v mezích, které jsou povolené Světovou zdravotnickou organizací, nevede k onemocnění štítné žlázy a dokonce může být její prevencí. Naopak lepší kvalita, ale zároveň i nižší obsah prospěšných prvků pro lidské zdraví může vést k onemocnění štítné žlázy.

V aplikaci byl k tomuto textu předložen pro názornost vodních nádrží pro jednotlivé obce Obr. 15 který zachycuje vodní nádrž zásobující Příbram.





**Obr. 15 Vodní nádrž Octárna**  
(Zdroj: Jiří Jiroušek 2018)

#### **4.2.5 Data od ČHMÚ a intenzita dopravy**

Data využívaná v práci, která se týkají znečištění ovzduší, jsou od Českého hydrometeorologického ústavu. Účelem ústavu je vykonávat funkci ústředního orgánu pro obory čistota ovzduší, hydrologie, jakost vody, klimatologie a meteorologie jako objektivní odborné služby poskytované přednostně pro státní správu České republiky. Hlavní činností je zřizovat a provozovat měřicí stanice a sítě s využíváním telekomunikačních sítí, odborně zpracovávat výsledky pozorování, měření a monitorování, dále poskytovat předpovědi a výstrahy a v neposlední řadě vytvářet a spravovat databáze (Portál ČHMÚ, 2016). Právě v těchto databázích lze nalézt velké množství dat a informací.

Od Českého hydrometeorologického ústavu (ČHMÚ) jsou volně k dispozici data, která se týkají znečištění ovzduší na jednotlivých měřicích stanicích za jednotlivé roky. V této práci jsou využita data za období let 2000 až 2005. Na internetových stránkách jsou poté k dispozici data o znečištění vnějšího ovzduší. Úroveň znečištění je způsobena vypouštěním znečišťujících látek z různých zdrojů v důsledku lidské činnosti, jako je doprava, průmyslová výroba a další. Data o znečištění vnějšího ovzduší jsou publikována v pětiletých průměrech a k dispozici jsou až od roku 2007. V práci jsou tedy využita data zprůměrovaná za rok 2007 až 2011.

K tomuto textu byl pro ilustraci přiložen Obr. 16 který znázorňuje komín teplárny Příbram a jeho úniky do ovzduší.



**Obr. 16 Komín teplárny Příbram**  
(Zdroj: Jiří Jiroušek 2018)

#### **4.2.5.1 Oxid siřičitý**

Oxid siřičitý je bezbarvý štiplavý plyn, nehořlavý a rozpustný ve vodě za vzniku kyselého roztoku. Jeho rozpustnost je silně závislá na teplotě.  $\text{SO}_2$  působí jako redukční činidlo, proto je využíván jako ochrana dřeva nebo při bělení. V potravinářství je využíván jako konzervační prostředek v alkoholových nápojích a sušeném ovoci. Místem nejvyššího využití oxidu siřičitého je průmysl výroby kyseliny sírové. Kapalný  $\text{SO}_2$  se v minulosti využíval k rafinaci ropných produktů a plyný jako ochranná atmosféra zabraňující oxidaci při tavení hořčíku, zde už byl ale nahrazen (IRZ 2018).

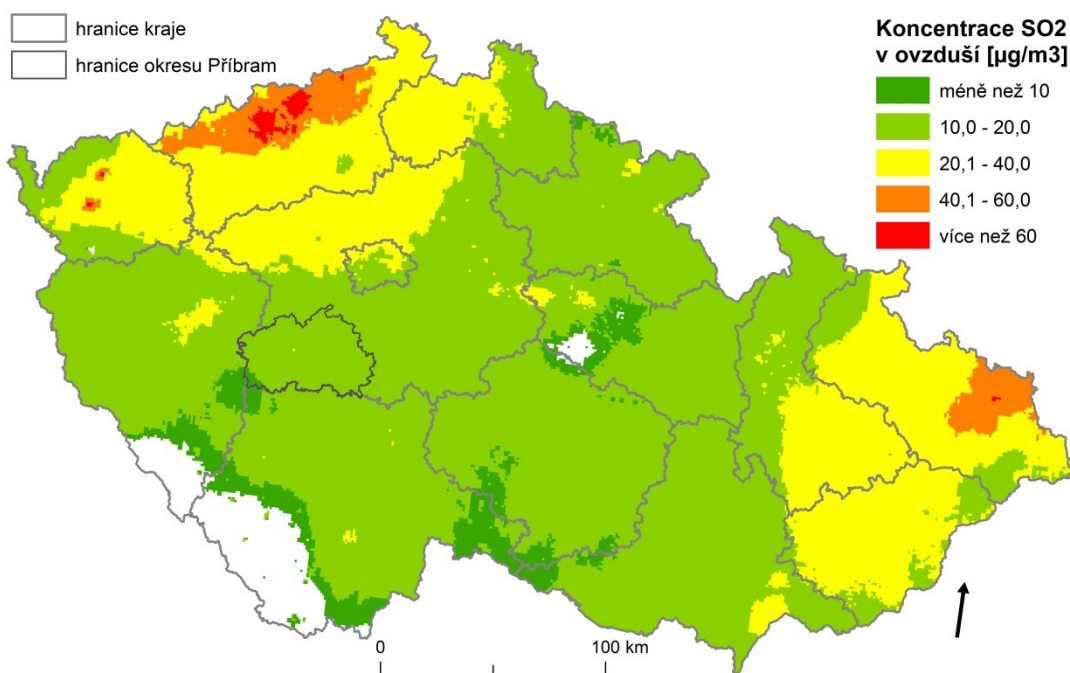
Mezi přírodní zdroje oxidů síry v ovzduší patří vulkanická činnost a přirozené lesní požáry. Lidským zdrojem je spalování paliv, kterého je v současné době velké množství. Jedná se o výrobu elektrické a tepelné energie, rafinerie ropy, dopravní prostředky a neméně významné zpracování kovů. Při všech těchto činnostech může docházet k oxidaci a následnému úniku do ovzduší. Při spalování tuhých paliv je cca 95 % přítomné síry přeměněno na  $\text{SO}_2$ , u kapalných paliv to může být až 100 % (IRZ 2018). Mnohdy je možné



použít účinná odsiřovací zařízení nebo jiné technologie, které mohou některé zdroje zcela omezit.

Oxid siřičitý může způsobovat velké množství negativních dopadů jak na životní prostředí, tak na zdraví člověka. Za určitý čas v ovzduší přechází fotochemickou nebo katalytickou reakcí na oxid sírový, který je hydratován vzdušnou vlhkostí na aerosol kyseliny sírové. Ta poté může reagovat s alkalickými částicemi prašného aerosolu za vzniku síranů. Síraný se dále usazují na zemský povrch nebo jsou z ovzduší vymývány srážkami. Při malém množství alkalických částic v ovzduší dochází k okyselení srážkových vod až na  $\text{pH} < 4$ . Takto oxidy síry společně s oxidy dusíku tvoří takzvané kyselé deště. Ty pak díky větru mohou být transportovány na velké vzdálenosti a způsobit značná poškození lesních porostů i průmyslových plodin, poškozují mikroorganismy, znehodnocují vodu, uvolňují z půdy kovové ionty a mohou způsobit úhyn ryb (IRZ 2018).

Při běžných koncentracích  $\text{SO}_2$  dráždí oči nebo horní cesty dýchací. Už při mírně vyšších koncentracích způsobuje zvýšení respirační nemocnosti u citlivých dospělých i dětí. Koncentrace  $0,5 \text{ mg} \cdot \text{m}^{-3}$  vede ke zvýšení úmrtnosti u starých chronicky nemocných lidí. Velmi ohroženou skupinou lidí jsou především astmatici, kteří bývají na působení oxidů síry velmi citliví. Při kontaktu s velmi vysokými koncentracemi oxidu siřičitého dochází k tvorbě tekutiny na plicích (IRZ 2018).



**Obr. 17** Průměrná koncentrace  $\text{SO}_2$  v ovzduší v letech 2007- 2011  
(Zdroj: vlastní)

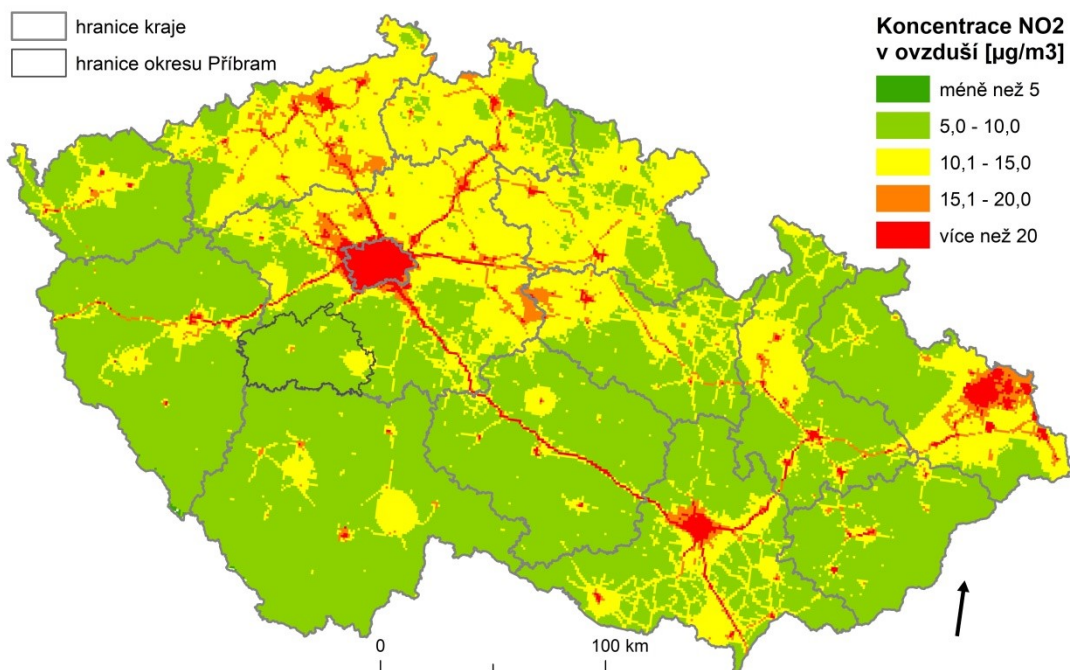
Povoleným imisním limitem je  $200 \mu\text{g}/\text{m}^3$ . Na území České republiky bylo v letech 2007 - 2011 dosaženo maximální cca.  $100 \mu\text{g}/\text{m}^3$ , a to na území Krušných hor. Druhou oblastí se zvýšenou koncentrací je Ostravsko. Na území okresu Příbram byla naměřena pouze mírná koncentrace, a to maximálně  $20 \mu\text{g}/\text{m}^3$ . Tato koncentrace byla naměřena ve městě Příbram. Vyšší koncentrace je poté na severu území. Celé území republiky je znázorněno na Obr. 17. V letech 2000 - 2005 byla naměřená koncentrace nižší a v severních okolních městech mimo okres byla koncentrace vyšší než v Příbrami.

#### **4.2.5.2 Oxid dusičitý**

Oxid dusičitý je červenohnědý plyn se štiplavým zápachem.  $\text{NO}_2$  je využíván jako meziprodukt ve výrobě kyseliny dusičné. Dalším použitím je silné meziokidačním činidlo v mnohých průmyslových procesech. Dále se vzhledem k jeho oxidačním vlastnostem může používat i v raketových palivech.

Emise oxidů dusíku jsou v současné době velmi úzce spojeny se spalováním i ušlechtilých paliv (plyn, nafta) a biomasy. Primární zdrojem jsou motorová vozidla, a to i přes užívání katalyzátorů. Při spalování ušlechtilých paliv v dopravních prostředcích a dalších motorových vozidlech je dosahováno vysoké teploty hoření. Proto zde dochází k oxidaci vzdušného dusíku na vysokoteplotní  $\text{NO}_x$ . V dnešní době mají tyto oxidy dusíku rostoucí charakter. Dalším zdrojem oxidů dusíku jsou veškeré chemické procesy, kdy jsou tyto oxidy přítomny a kde může dojít k jejich úniku. Tyto výroby jsou ale dnes často vybaveny účinným zařízením na odstranění oxidů dusíku z odpadních plynů. Mezi přírodní zdroje oxidů dusíku patří oxidace v půdách nebo oxidace vzdušného dusíku během výbojů v atmosféře (IRZ 2018).

Na zdraví člověka mohou oxidy dusíku působit negativně hlavně ve vyšších koncentracích, které nejsou ale běžně v ovzduší přítomny. Vdechování vysokých koncentrací vede k závažným zdravotním potížím a může dokonce způsobit smrt (IRZ 2018). Některé studie ukazují, že oxidy dusíku mají určitou roli i při vzniku nádorových onemocnění.



**Obr. 18 Průměrná koncentrace NO<sub>2</sub> v ovzduší v letech 2007 – 2011**  
(Zdroj: vlastní)

Na přiložené mapě (Obr. 18) znázorňující průměrnou koncentraci NO<sub>2</sub> v ovzduší v letech 2007 - 2011 jsou výrazně vidět tři největší česká města (Praha, Brno a Ostrava). Dále jsou zde výrazné velké dopravní tahy pro motorová vozidla, a to hlavně dálnice D1, D2, D5, D8 a D10.

### Doprava

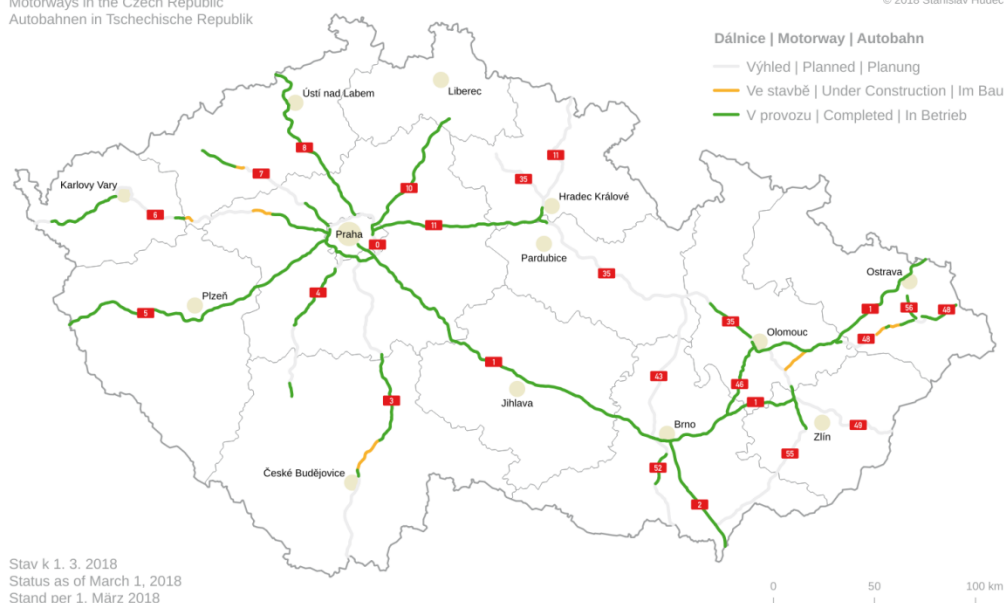
V okrese Příbram je taktéž viditelný nejdůležitější dopravní tah, a to dálnice D4. Dále jsou zde velmi vidět Sedlčany a jejich blízké okolí, které jsou vidět i z mapy celé ČR, jako oblast se zvýšenou koncentrací. Celá dálniční síť je znázorněna na Obr. 19. Pravděpodobně na to bude mít vliv blízkost k dálnici D3, která je hlavním tahem z Prahy na jih Čech.

## Dálnice v České republice

Motorways in the Czech Republic  
Autobahnen in Tschechische Republik

ceskedalnice.cz

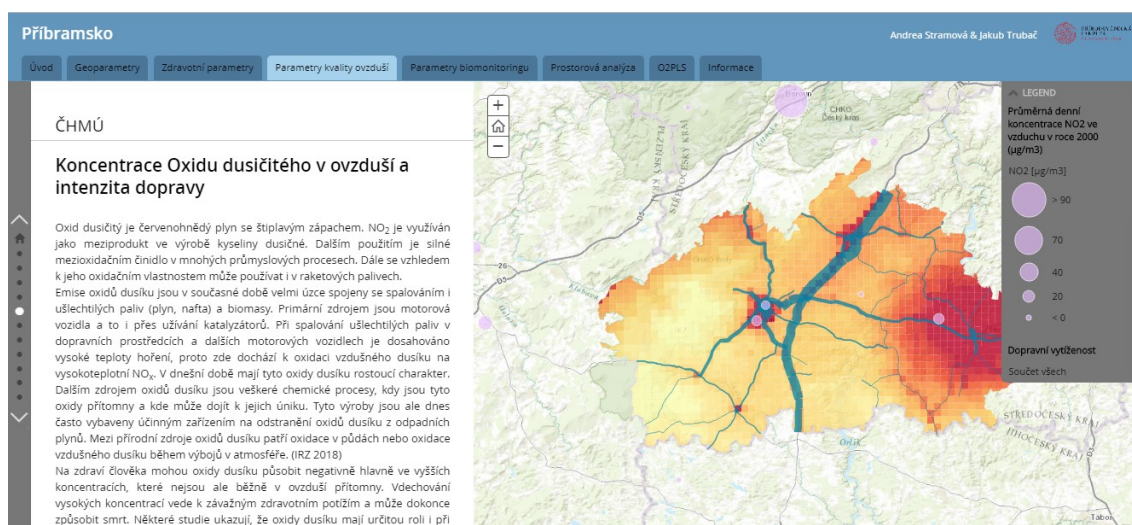
© 2018 Stanislav Hudec



**Obr. 19 Dálnice v České republice.**

(Zdroj: ceskedalnice.cz 2016)

Na Obr. 20 je snímek prezentace předkládaných informací pomocí webové aplikace.



**Obr. 20 Prezentace parametru NO<sub>2</sub> pomocí webové aplikace**

### 4.2.5.3 Zhodnocení vlivu parametrů emisí na funkci štítné žlázy

Od ČHMÚ prezentována data za dvě období. První období je od roku 2000 do roku 2005 a jedná se o bodovou vrstvu, která znázorňuje koncentraci 8 prvků v ovzduší. Jedná se o oxid siřičitý, prašný aerosol, oxidy dusíku, mangan, nikl, olovo, arsen a částice s velikostí frakce po 10 µm. Tyto parametry byly vybrány, protože byly měřeny na nejvíce stanicích. Každá měřicí stanice měří pouze určité vybrané parametry a tyto byly měřeny nejčastěji na 23 vybraných

stanicích v okrese Příbram a okolí. U tří měřených parametrů (Mn, Ni, Pb) byla nejvyšší koncentrace naměřena v Příbrami. Zbylé nejvyšší koncentrace byly naměřeny v Plzni nebo Benešově. Benešov je o 20 km blíž k okresu Příbram než Plzeň a zároveň díky západnímu proudění větru má Benešov větší vliv na okres Příbram. To je potvrzeno i v polygonové vrstvě, která reprezentuje průměrnou koncentraci parametrů v letech 2007-2011. V této vrstvě je znázorněno 11 parametrů. Ve většině vrstev je vidět, že nejvyšší koncentrace je vždy dosažena v Příbrami a v severovýchodní části okresu, mimo koncentrace olova a kadmia, která je nejvyšší v největších obcích okresu.

Při srovnání parametrů funkce štítné žlázy s parametry ovzduší není přímo viditelná souvislost. Předpokladem by bylo, že v Příbrami a severovýchodní části okresu budou bydlet pacienti s větším počtem překročených norem, tedy poruchou funkce štítné žlázy. Tato hypotéza se přímo z map nedá potvrdit; proto je nutné přistoupit ke statistickému popisu zhoršené funkce štítné žlázy, který je popsán v diplomové práci Stramové (2017), a poté k prostorové analýze endokrinologických dat a dat mapující životní prostředí v okrese Příbram. Tyto prostorové analýzy jsou znázorněny v následující části a zároveň popsány v diplomové práci Stramové (2017).

Pro názornost byl k tomuto textu ve webové aplikaci přiložen Obr. 21 který zaznamenává znečištěné ovzduší v okolí těžebních věží.



**Obr. 21 Ukázka znečištěného ovzduší**  
(Zdroj: Tomáš Hájek, Lidové noviny 2011)



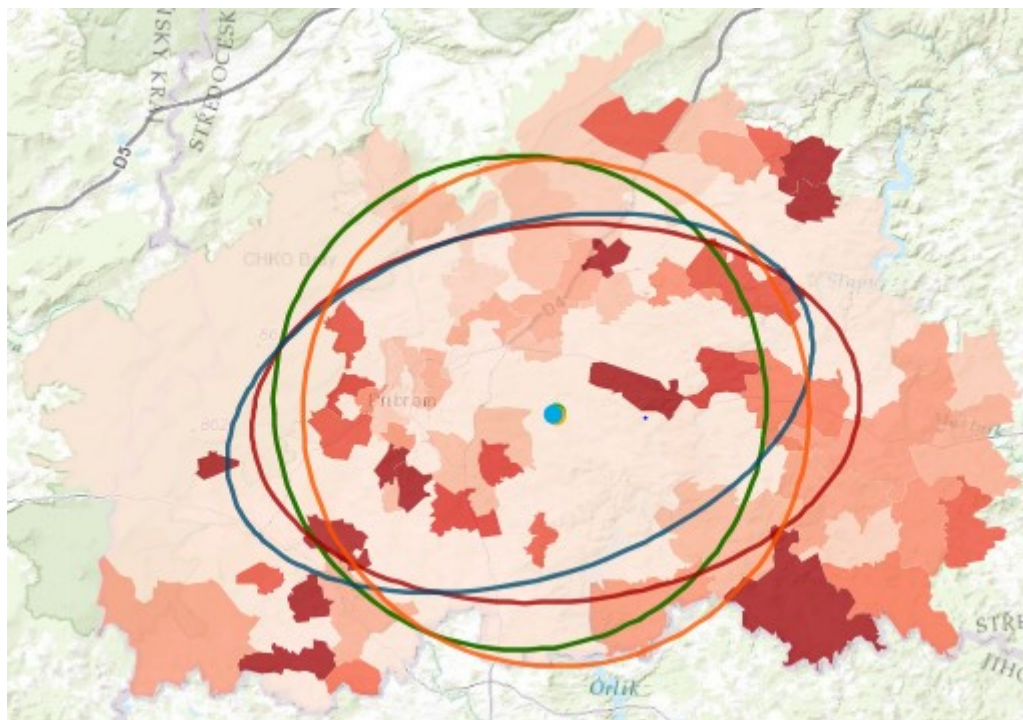
## 4.3 Prostorová analýza

### 4.3.1 Analýza prostorových vzorů

V této mapě (Obr. 22) byly vizualizovány výsledky popisné prostorové statistiky. Geografický střed je průměr souřadnic, který může být vážený a nevážený. Geografický střed všech vyšetřených pacientů se nachází v severní části obce Jablonná, která leží ve středu okresu. Geografický střed vážený počtem překročených norem je posunut mírně jihovýchodně od geografického středu neváženého, ale stále zůstává v obci Jablonná. Velice blízko se nachází geografický střed pouze pro pacienty s překročenou normou 2 a vícekrát. Na základě geografického středu je zřejmé, že pacienti s překročenými normami jsou rozmístěni po celém okresu s výjimkou CHKO Brdy, kde nežijí žádní obyvatelé. Tuto neobydlenost vyrovnává největší město okresu Příbram, které leží v těsné blízkosti CHKO Brdy. Zbýlá města Dobříš, Březnice a Sedlčany se díky svému počtu obyvatel podílejí na tom, že všechny geografické středy leží přibližně uprostřed území.

Směrodatná vzdálenost je míra rozptylu hodnot kolem geografického středu. Jedná se o kružnici kolem geografického středu, jejímž poloměrem je hodnota směrodatné vzdálenosti (Netrdová [online]). U všech pacientů je směrodatná vzdálenost 15 229,5 m, u pacientů s překročenou normou 2 a více je vypočítaná hodnota 15 650,1 m. Rozdíl mezi směrodatnými vzdálenostmi je přibližně 400 m. U všech vyšetřených pacientů je rozptyl kolem geografického středu menší než u pacientů s překročenou normou 2 a vícekrát. Tento výsledek potvrzuje předchozí zjištění z výsledků zkoumání polohy geografického středu, tedy ten, že pacienti jsou rozmístěni po celém okrese (Stramová 2017).

Směrodatná elipsa odchylek vykazuje, zda prostorové rozmístění jevů vykazuje rysy směrovosti (Netrdová [online]). Směrodatná elipsa všech vyšetřených pacientů je užší a mírně protáhlejší oproti směrodatné elipse pacientů s překročenou normou 2 a vícekrát. Nejvýraznější rozdíl je v jejich směru. Směr elipsy, znázorňující pacienty s překročenou normou 2 a vícekrát, je posunut přibližně o 10 ° protisměru hodinových ručiček. Tento výsledek naznačuje rozmístění pacientů s překročenou normou 2 a více hlavně v oblasti měst Příbram a Sedlčany.



**Obr. 22 Analýza prostorových vzorů**  
(Zdroj: Vlastní)

Polygonová vrstva, která je vizualizována v mapě, znázorňuje počet překročených norem na počet obyvatel v dané obci. Tato vrstva byla vytvořena, aby eliminovala vliv většího počtu obyvatel ve městě a bylo možné mezi sebou jednotlivé obce porovnat. Město Příbram je zde dokonce rozděleno na jednotlivá katastrální území. Největší překročení norem na 1000 obyvatel bylo zjištěno v obci Hudčice (22,4) a v obci Hlubyně (22,1). Dohromady bylo u 9 obcí zjištěno 10 a více překročení norem na 1000 obyvatel. Města Příbram, Dobříš a Sedlčany zde nejsou nijak výrazná, počet překročených norem na 1000 obyvatel je do 5.

V diplomové práci Stramové (2017) je podrobně rozebrána prostorová analýza těchto dat. Práce zaznamenává celou konfirmační analýzu prostorové statistiky. Jsou zde vyobrazeny výsledky koncentrace počtu překročených norem na  $1 \text{ km}^2$  jak pro všechny pacienty, tak pro jednotlivé ukazatele funkce štítné žlázy. Stejně tak je vytvořen i počet překročených norem na 1000 obyvatel. Dále je zde aplikována metoda nejbližšího souseda, jejímž výsledkem je, že pacienti jsou rozmístěni ve shlucích, které nejsou náhodné, a to na hladině pravděpodobnosti 99 %. Tento výsledek potvrdila i druhá metoda, a to Ripleyho K- funkce. Poslední použitá metoda Lokální Moranovo I identifikovala statisticky významná seskupení překročených norem pro jednotlivé parametry funkce štítné žlázy.

### 4.3.2 Variogramy endokrinologických parametrů

K vytvoření jednotlivých grafů byl využit program ArcMap 10.3.1 s licencí pro Geostatistical Analyst. Nejdříve byl u všech dat proveden výběr, následně byl vytvořen histogram pro analýzu rozložení hodnot v datovém souboru. Dalším krokem byla analýza trendu a vytvoření variogramu mrak. V posledním kroku byl vytvořen variogram, odhalující nepodobnost v datech.

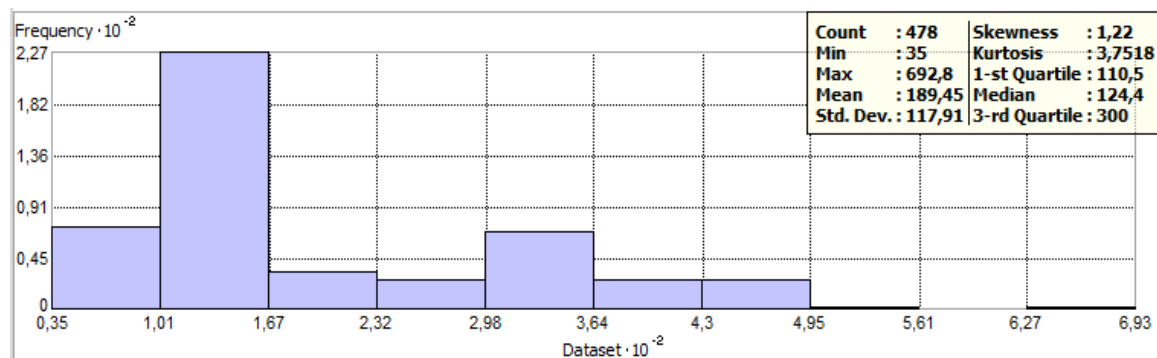
#### 4.3.2.1 Obsah jódu v moči

U ukazatele Jodurie muselo být odstraněno 12 pacientů, u kterých nebyla jodurie měřena. Analýza s variogramem je založena na 478 pacientech. Prvně byly zjištěny základní charakteristiky, které jsou vypsány v Tab. 10. Nejmenší hodnota jodurie je 35  $\mu\text{g/l}$  a nejvyšší 692,8  $\mu\text{g/l}$ . Při správné funkci štítné žlázy je naměřená hodnota obsahu jódu v moči 100 až 300  $\mu\text{g/l}$ . Průměrná hodnota, stejně jako medián, odpovídá správnému fungování štítné žlázy, ale jsou výrazně nižší než maximum, spíše se blíží k minimu. To naznačuje, že převažují nižší hodnoty. Toto potvrzuje i hodnota šikmosti 1,22. Směrodatná odchylka je zde poměrně výrazná a říká, že až cca 68 % dat se pohybuje v rozmezí  $\pm 117,91$  od průměru.

Ukazatel	Hodnota
Minimum	35
Maximum	692,8
Průměr	189,45
Medián	124,4
První kvartil	110,5
Třetí kvartil	300
Směrodatná odchylka	117,91
Šikmost	1,22

Tab. 10. Základní ukazatele rozložení hodnot u Jodurie

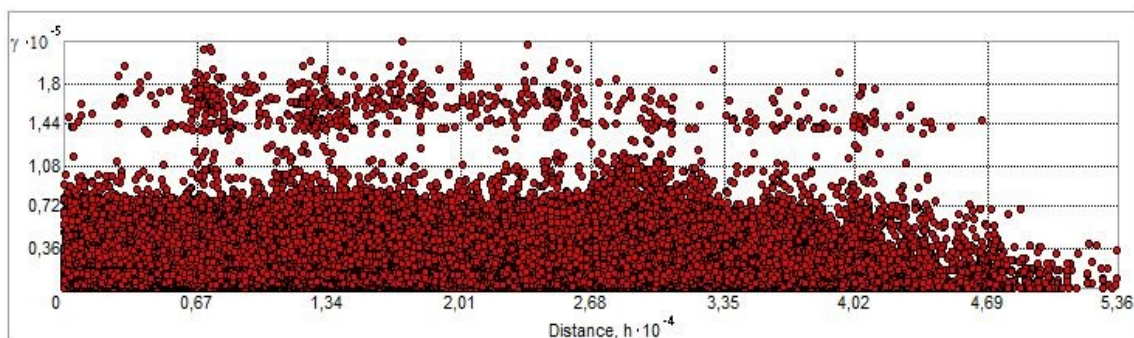
Následně byl vytvořen histogram (Graf 2). Z něj je vidět nesymetrické rozložení hodnot Jodurie. Zároveň je potvrzena hodnota šikmosti, která dokazovala, že více hodnot se nalézá vlevo a vpravo od průměru se vyskytují odlehlejší hodnoty.



Graf 2. Histogram naměřených hodnot Jodurie

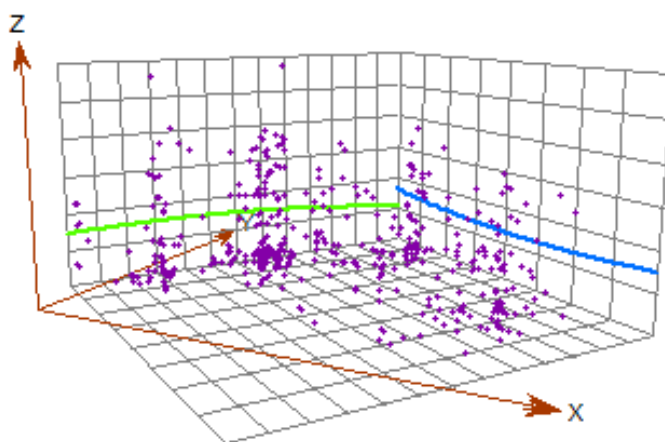


Variogram mrak (Graf 3) pro dvojice bodů u hodnot jodurie je zobrazen na následujícím grafu. Na ose Y je zaznamenána míra nepodobnosti. Zde je vidět předěl mezi hlavní skupinou a méně početnou skupinou s velkou mírou nepodobnosti. Rozdělení hodnot do dvou skupin není patrné hned od počátku. Významné se stává až cca od 700 m.



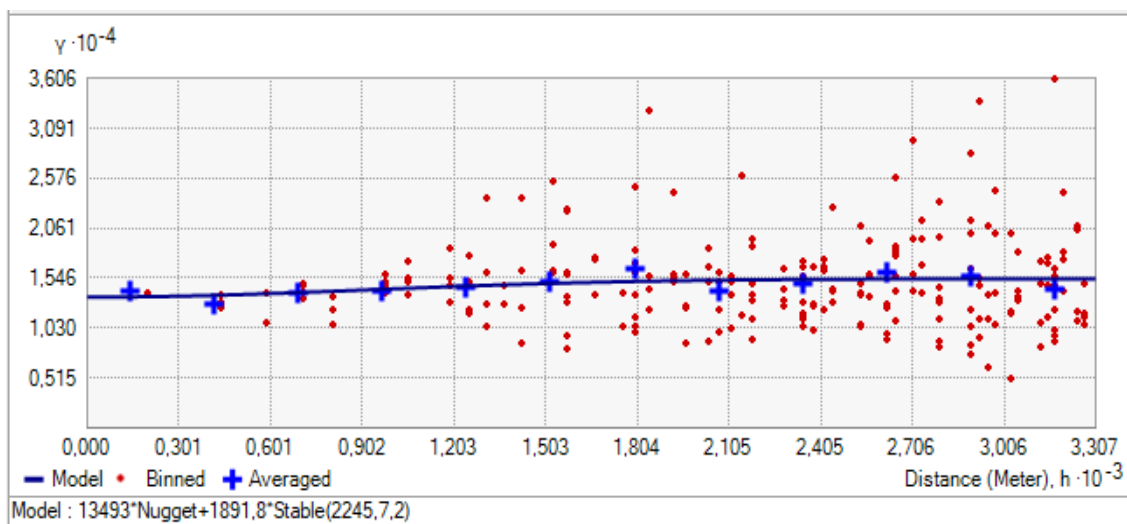
**Graf 3.** Variogram mrak naměřených hodnot Jodurie

U analýzy trendu (Graf 4) nevykazuje nijak výrazný trend. Fialové body představují jednotlivé naměřené hodnoty Jodurie.



**Graf 4.** Analýza trendu u naměřených hodnot Jodurie

V posledním kroku u dat Jodurie byl vytvořen variogram (Graf 5), který je zobrazen na následujícím grafu. Pro jeho tvorbu byly využity všechny naměřené hodnoty. Velikost kroku je 276 m, graf je rozdělen na 12 tříd. Míra korelace je zde velice malá a navíc pouze na krátké vzdálenosti cca do 150 m. Tato korelace bude pravděpodobně velice slabá a možná i neexistující.



**Graf 5. Variogram pro naměřené hodnoty Jodurie**

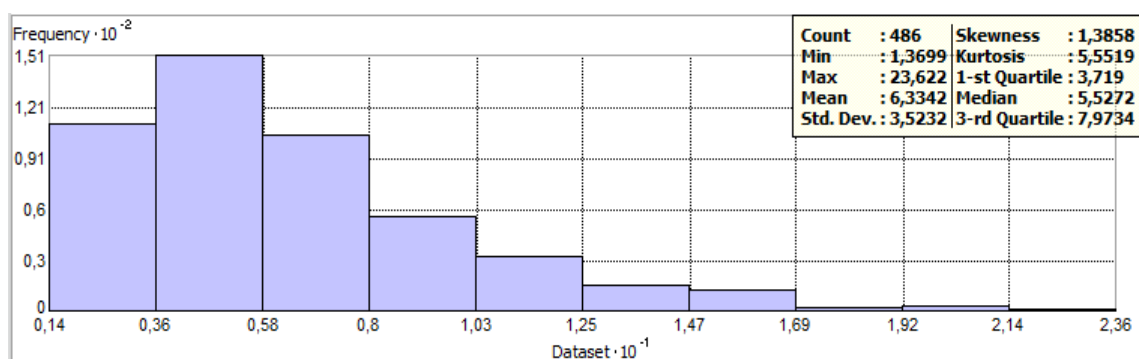
#### 4.3.2.2 Objem štítné žlázy

Z celkového počtu 490 vyšetřených pacientů nebyli do analýzy pro objem štítné žlázy zahrnuti čtyři pacienti, jelikož u nich nebyly změřeny hodnoty Volgut. Celkově tedy analýza dat a variogram byly vytvořeny pro 486 pacientů. Nejprve byly zjištěny základní charakteristiky, které jsou zaznamenány v Tab. 11. Medián i průměr jsou výrazně nižší než maximum a blíží se k minimu. Toto dokazuje i koeficient šikmosti, který je kladný, což znamená, že vpravo od průměru se vyskytují odlehlejší hodnoty nežli vlevo. Rozdělení má tzv. pravý ocas, což znamená, že většina hodnot se nalézá vlevo od průměru. Směrodatná odchylka je v tomto souboru dat poměrně velká a interpretuje, že by až cca 68 % dat mělo být od průměru vzdáleno do hodnoty 3,5232.

Ukazatel	Hodnota
Minimum	1,36994
Maximum	23,6224
Průměr	6,2825
Medián	5,5272
První kvartil	3,719
Třetí kvartil	7,9734
Směrodatná odchylka	3,5232
Šikmost	1,3858

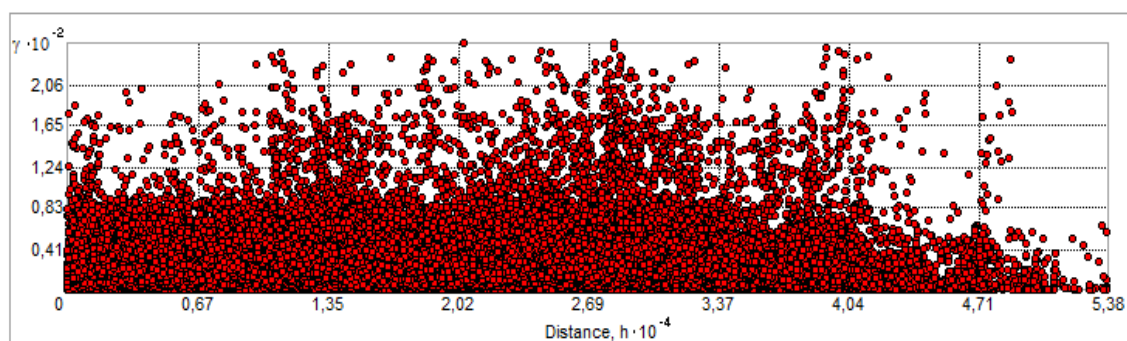
**Tab. 11. Základní ukazatele rozložení hodnot u Volgut**

V programu ArcMap 10.3.1 byl vytvořen histogram s rozdělením dat (Graf 6). Z histogramu je vidět, že rozdělení dat není symetrické a nejvíce hodnot se nalézá v první třetině grafu.



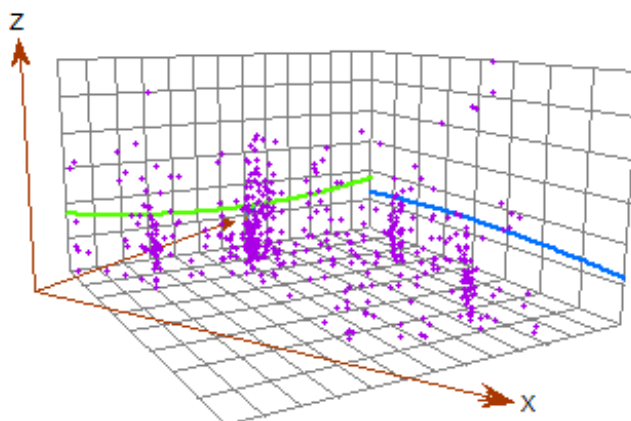
**Graf 6. Histogram naměřených hodnot Volgut**

Následně byl vytvořen v programu ArcMap 13.3.1 variogram mrak pro všechny dvojice bodů (Graf 7). Z následujícího grafu je vidět, že míra variability v jedné lokalitě je přibližně stejná jako míra variability v celém prostoru. To by mohlo značit, že velikost objemu štítné žlázy se mění nezávisle.



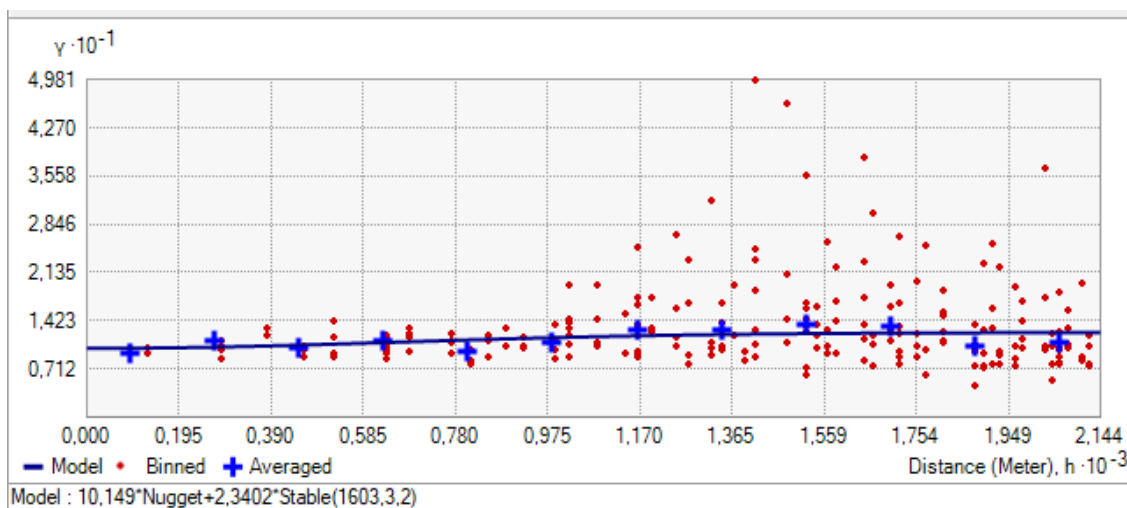
**Graf 7. Variogram mrak naměřených hodnot Volgut**

Z analýzy trendu (Graf 8), která je zobrazena na následujícím grafu, také není zřetelný trend.



**Graf 8. Analýza trendu u naměřených hodnot Volgut**

V poslední části byl vytvořen samotný variogram (Graf 9). Při tvorbě byly využity všechny naměřené hodnoty. Velikost kroku byla zvolena 179 m s rozdělením na 12 tříd. Zde je možné vidět velmi malou míru korelace, a to ještě na velmi krátké vzdálenosti do přibližně 100 metrů. Tato korelace je pravděpodobně velmi slabá.



**Graf 9.** Variogram pro naměřené hodnoty Volgut

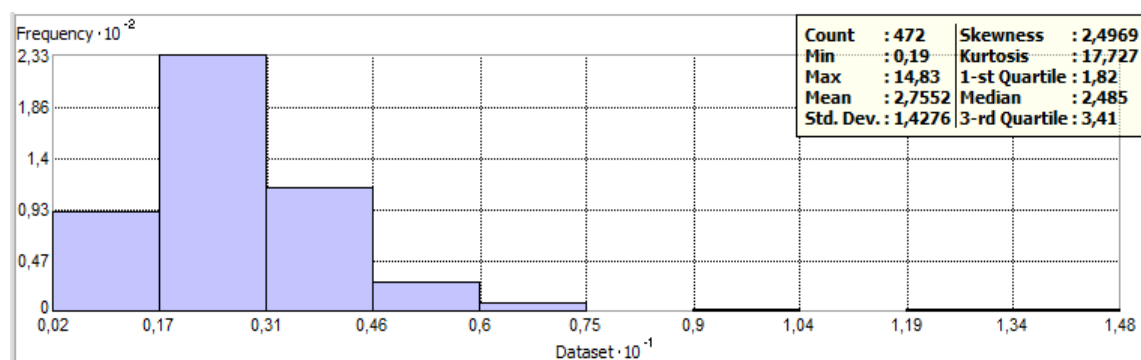
#### 4.3.2.3 TSH

I u těchto dat museli být nejdříve z analýzy vyjmuty pacienti bez změřených hodnot. Celkově byla analýza u hormonu TSH provedena ze souboru 472 pacientů. Minimální hodnota TSH byla 0,19 mIU/l a maximální 14,83 mIU/l. Při správném fungování štítné žlázy je hladina tohoto hormonu od 0,27 do 4,2 mIU/l. V tomto rozmezí se pohybují hodnoty průměru i mediánu, které jsou si podobné. Stejně jako u Jodurie a Volgut i zde má průměr a medián blíže k minimu, což ukazuje i koeficient šikmosti s hodnotou 2,4969. To vypovídá o tom, že více odlehklých hodnot leží od průměru vpravo. Směrodatná odchylka je 1,4276.

Ukazatel	Hodnota
Minimum	0,19
Maximum	14,83
Průměr	2,76
Medián	2,49
První kvartil	1,82
Třetí kvartil	3,41
Směrodatná odchylka	1,43
Šikmost	2,49

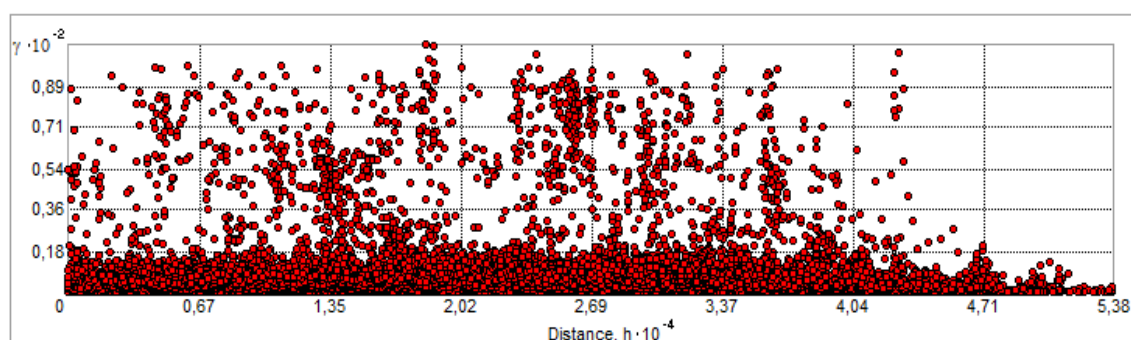
**Tab. 12.** Základní ukazatele rozložení hodnot TSH

Vytvořený histogram (Graf 10) by měl pravidelné rozložení, kdyby nebylo tolik odlehlých měření, která mají vysokou hodnotu TSH. Toto tvrzení přesně potvrzuje koeficient šikmosti, který je u těchto dat opravdu veliký.



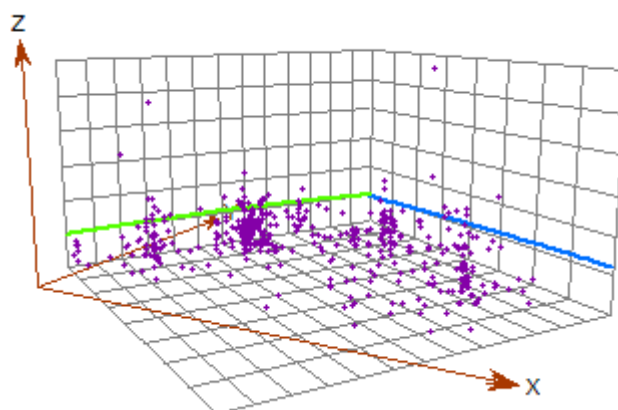
**Graf 10. Histogram naměřených hodnot TSH**

Variogram mrak (Graf 11) pro hodnoty TSH nemá tak pravidelné rozložení na ose Y u míry nepodobnosti. Většina hodnot se nachází spíše v okolí nízkých hodnot míry nepodobnosti. Z grafu ale není patrný nějaký trend v prostorové závislosti.



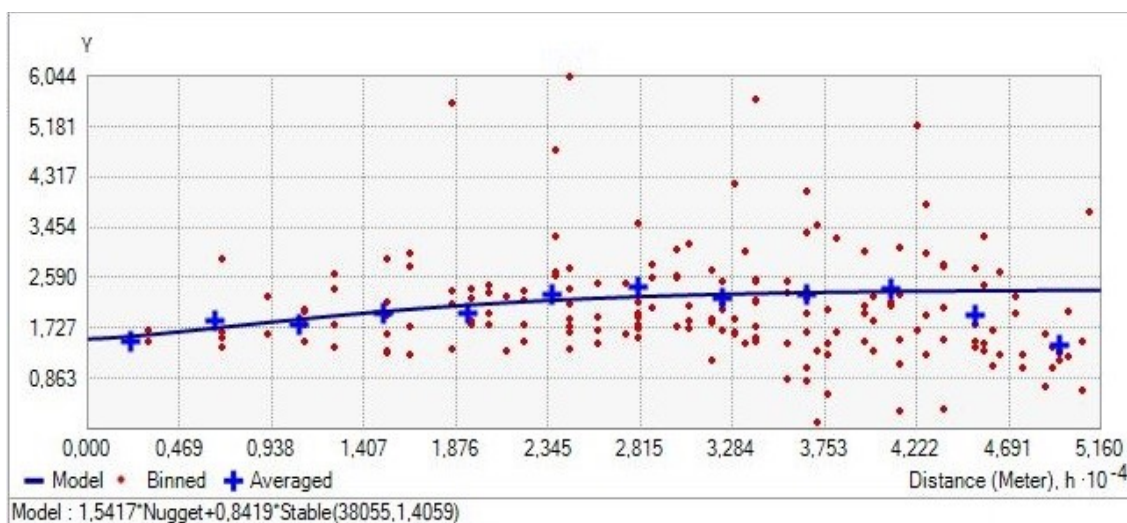
**Graf 11. Variogram mrak naměřených hodnot TSH**

Na grafu analýzy trendu (Graf 12) nebyl nalezen na první pohled žádný trend.



**Graf 12. Analýza trendu u naměřených hodnot TSH**

Stejně jako při analýze trendu ani zde není jasně patrná korelace jednotlivých bodů. Prostorová struktura je slabá, ale je vidět, že variogram nějakou závislost naznačuje (Graf 13). Krokem variogramu je 4300 m a počet tříd je stejně jako u předchozích variogramů 12.



**Graf 13.** Variogram pro naměřené hodnoty TSH

#### 4.3.2.4 FT3

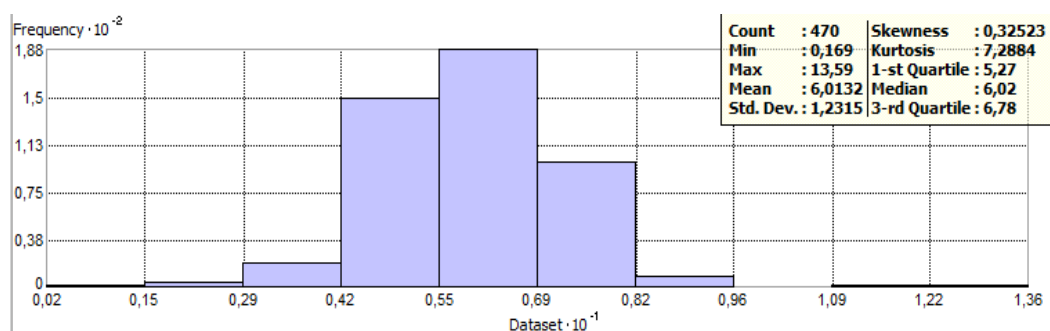
Stejně tak jako u předchozích ukazatelů funkce štítné žlázy i zde bylo nutné z původního datového souboru vybrat pouze pacienty, u kterých byla hodnota FT3 změřena. Těchto pacientů bylo 470. Minimální naměřená hodnota je 0,169 pmol/l a maximální je 13,59 pmol/l. Při normální funkci štítné žlázy by hodnota FT3 měla dosahovat od 2,8 do 7,1 pmol/l. Průměr i medián se do těchto mezí vejdou. Díky ukazateli šikmosti, který je 0,32523, lze říct, že rozložení dat by mělo být přibližně symetrické s malým „ocasem“ odlehlých hodnot vpravo. Směrodatná odchylka je zde oproti předchozím datovým souborům menší (Tab. 13).

Ukazatel	Hodnota
Minimum	0,17
Maximum	13,59
Průměr	6,01
Medián	6,02
První kvartil	5,27
Třetí kvartil	6,78
Směrodatná odchylka	1,23
Šikmost	0,33

**Tab. 13.** Základní ukazatele rozložení hodnot FT3

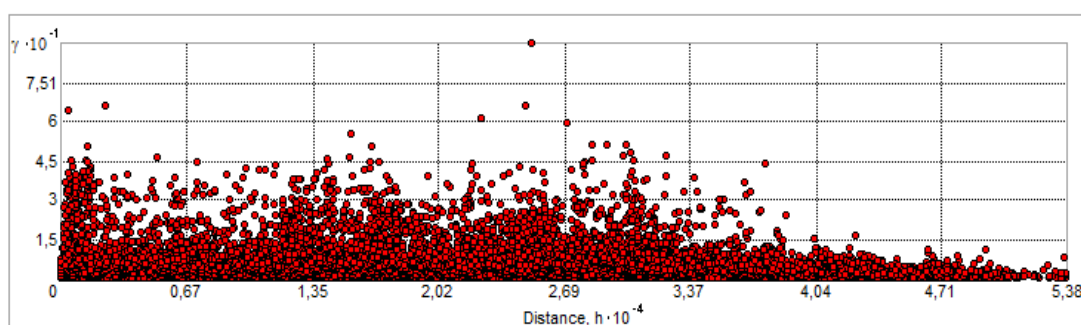


U histogramu (Graf 14) je vidět symetrické rozložení, kdy nejvíc hodnot se nalézají v rozmezích normální produkce hormonu FT3. Zároveň jsou zde vidět odlehle hodnoty vpravo, čímž se potvrzuje pouze mírně kladný koeficient šikmosti.



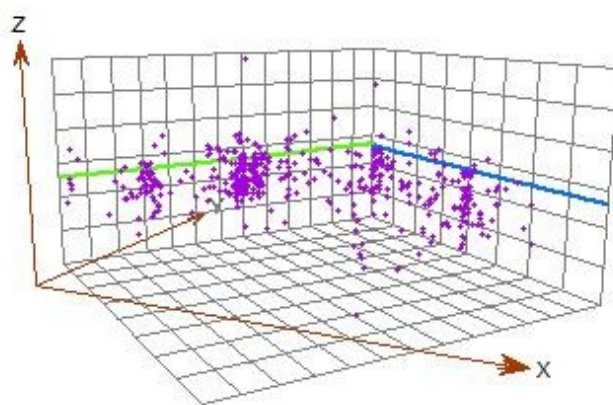
**Graf 14. Histogram naměřených hodnot FT3**

Variogram mrak (Graf 15) pro naměřené hodnoty hormonu FT3 neukazuje téměř žádnou prostorovou závislost. Naopak se zvyšující se vzdáleností ubývá hodnot, které mají větší míru nepodobnosti.



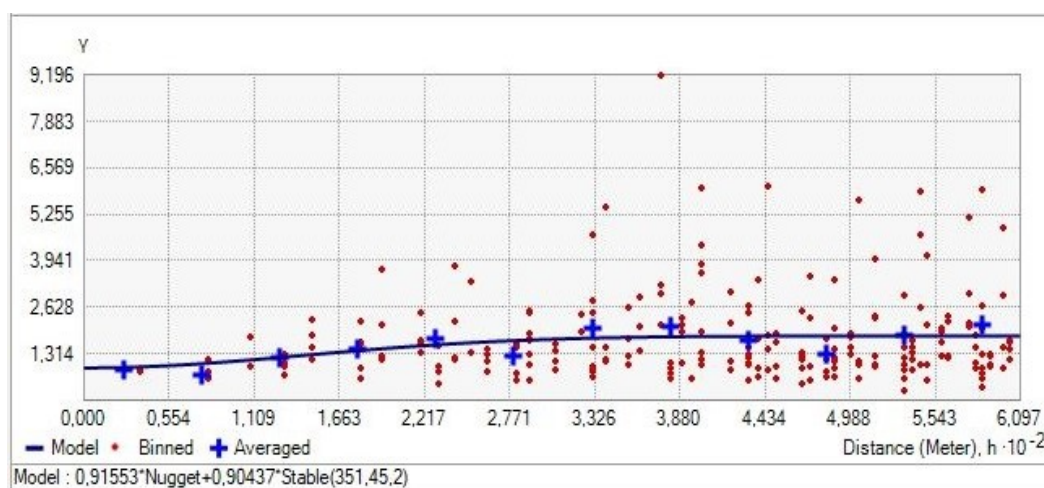
**Graf 15. Variogram mrak naměřených hodnot FT3**

Z grafu analýzy trendu (Graf 16) není patrný žádný trend anebo je pouze lokální a na grafu se neprojevuje. Viditelné je, že rozsah hodnot je podobný jak na lokální úrovni, tak na úrovni měst i celého okresu.



**Graf 16. Analýza trendu u naměřených hodnot FT3**

Variogram (Graf 17) pro naměřené hodnoty hormonů FT3 byl vytvořen s krokem 50 m a s počtem tříd 12. Je zde vidět možná prostorová závislost na malé vzdálenosti cca 200 m. Hodnota nugget efektu je přibližně 0,9. Podobná je i hodnota pro práh. To znamená, že zhruba 50 % měření by mohlo být modelem variogramu vysvětleno, ale pouze na velmi krátké vzdálenosti (max. 100 m). Protože je tato prostorová struktura slabá, není možné díky ní predikovat hodnoty na delší vzdálenosti a na krátké vzdálenosti jen maximálně z 50 %.



**Graf 17. Variogram pro naměřené hodnoty FT3**

#### 4.3.2.5 FT4

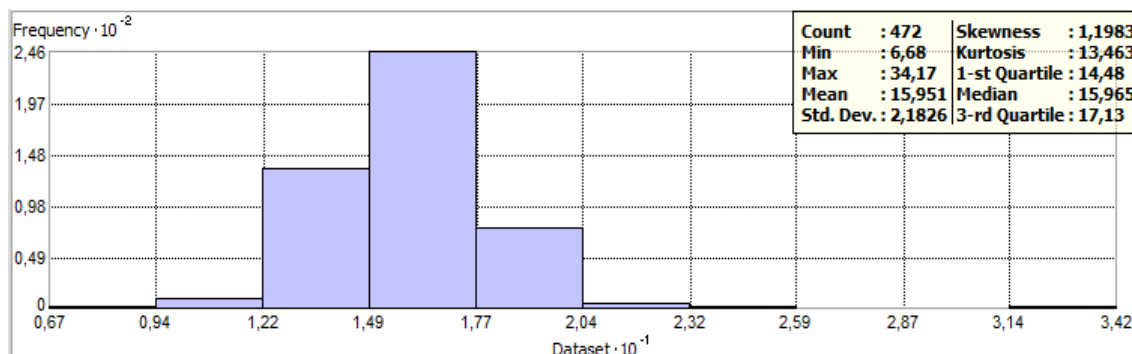
Rovněž u posledního analyzovaného parametru funkce štítné žlázy byly nejprve odstraněny neměřené hodnoty. Pro analýzu zůstalo 472 naměřených hodnot. Minimální naměřenou hodnotou 6,68 pmol/l a maximální naměřenou hodnotou bylo 34,17 pmol/l. Spodní hranicí hormonu pro normální funkci štítné žlázy je hodnota 12 pmol/l a horní hranicí je 24 pmol/l. Do těchto mezí se vejde hodnota průměru i mediánu. Tyto dvě hodnoty jsou si velmi podobné a liší se až v řádu setin. Obě hodnoty mají blíže k minimu než k maximu. Směrodatná odchylka je 2,1826 a hodnota šikmosti 1,1983, což značí více odlehlých hodnot vpravo od průměru.

Ukazatel	Hodnota
Minimum	6,68
Maximum	34,17
Průměr	15,95
Medián	15,97
První kvartil	14,48
Třetí kvartil	17,13
Směrodatná odchylka	2,18
Šikmost	1,19

**Tab. 14. Základní ukazatele rozložení hodnot FT4**

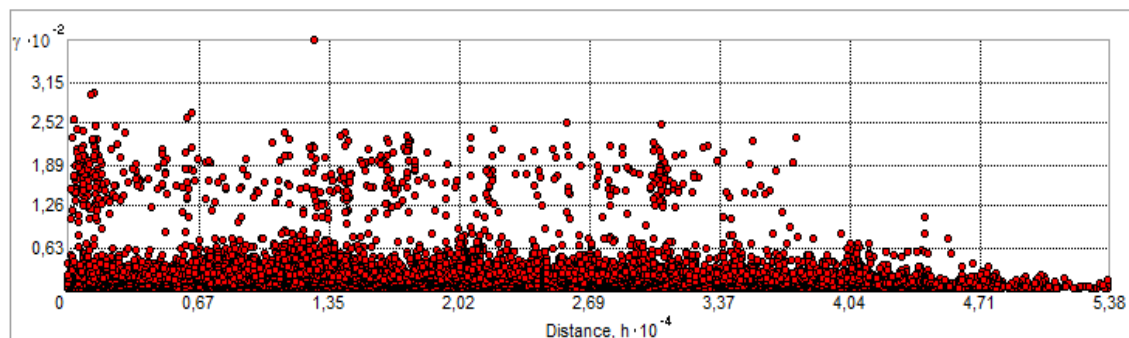


Pro potvrzení hodnoty šikmosti byl vytvořen histogram (Graf 18). Z něj jsou vpravo odlehle hodnoty patrné. Kdyby nebylo těchto hodnot, je rozložení naměřených hodnot symetrické.



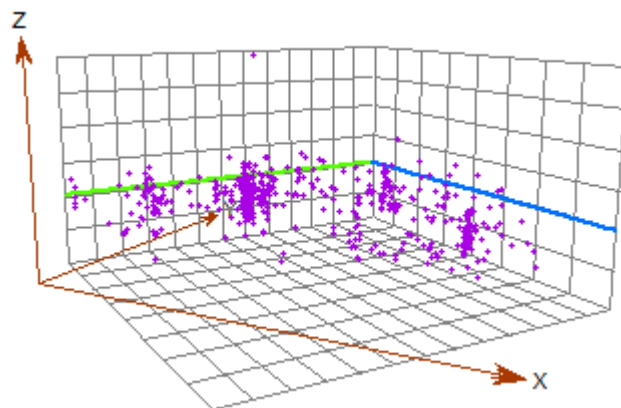
**Graf 18. Histogram naměřených hodnot FT4**

Variogram mrak (Graf 19) pro FT4 je podobný variogramu mrak pro hodnoty hormonu FT3. Zde je ale výrazný předěl mezi podobnými hodnotami a pak hodnotami nepodobnými, který není vyplněn přechodovými hodnotami. S narůstající vzdáleností se tento předěl ztrácí a hodnoty jsou si více podobné.



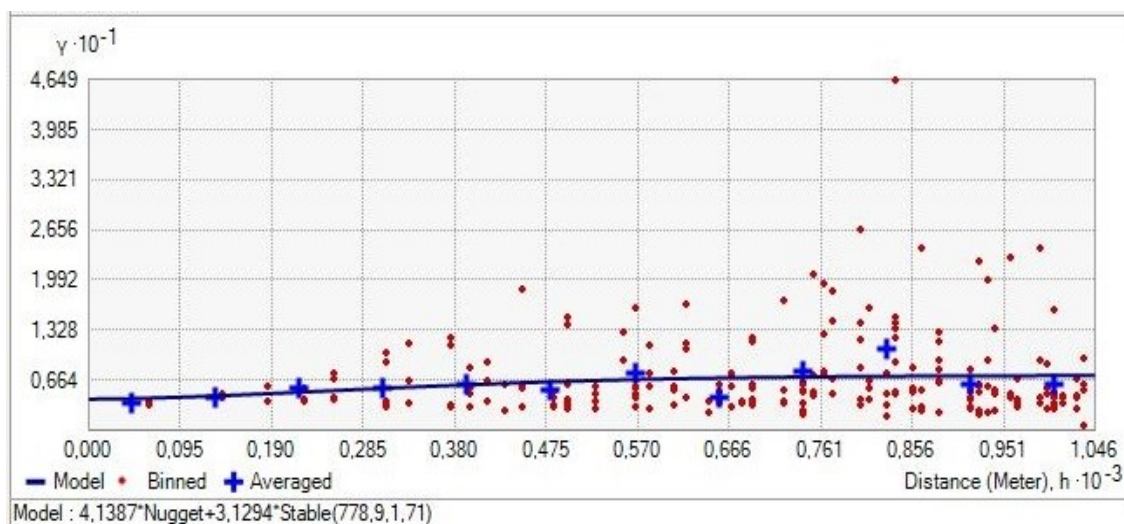
**Graf 19. Variogram mrak naměřených hodnot FT4**

Ani u hormonu FT4 není z grafu analýzy trendu (Graf 20) nijak patrný trend. Stejně jako u hormonu TSH a FT3 nemají naměřené hodnoty mezi sebou tak velké rozdíly. U hodnot Jodurie a Volgotu bylo možné sledovat větší rozdíly naměřených hodnot v obcích.



**Graf 20. Analýza trendu u naměřených hodnot FT4**

Do tvorby jsou vždy zahrnuty veškeré naměřené hodnoty i přesto, že jejich prostorové umístění si může být velice blízké. Pro tvorbu variogramu byl zvolen krok 90 a počet tříd 12. Variogram (Graf 21) ukazuje naprosto minimální prostorovou závislost. Dokonce je možné, že zde žádná prostorová závislost neexistuje.

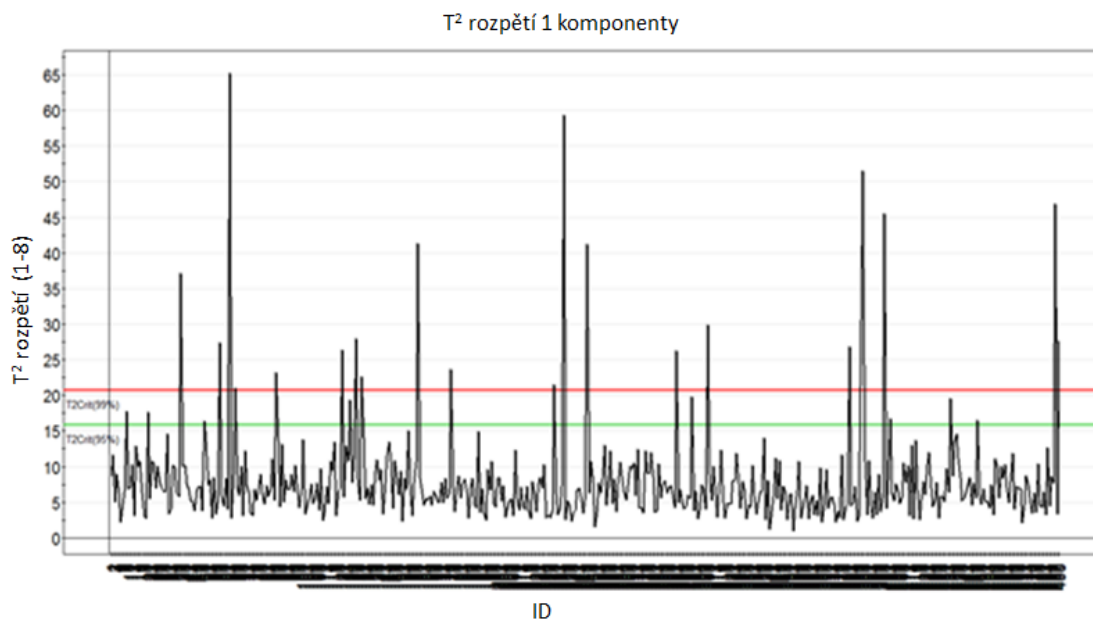


**Graf 21. Variogram pro naměřené hodnoty FT4**

## 4.4 O2PLS/OPLS

Pro analýzu vztahů mezi přítomností nemoci a vnějšími faktory byla použita vícerozměrná regrese s redukcí dimenzionality, metoda obousměrných ortogonálních projekcí do latentní struktury (O2PLS) nebo její jednodušší varianta OPLS a to v případech kdy byla zvolena pouze jedna závisle proměnná v modelu. V této práci řešíme otázku vztahů mezi dvěma skupinami proměnných (lékařsko-diagnostická data vs. charakteristiky životního prostředí). Metody PCA a FA nejsou optimální z hlediska kladené otázky. Obě metody by bylo možné aplikovat, ale výsledek by byl hůře interpretovatelný a méně adekvátní, než v případě O2PLS. Vysvětlení, proč je pro daný účel tato metoda vhodnější než analýza hlavních komponent a faktorová analýza, se nachází v diplomové práci Stramové (2017). Cílem O2PLS není interpretace vztahů uvnitř matice prediktorů  $X$  či predikovaných proměnných  $Y$ , i když tyto vztahy jsou vysvětlovány v ortogonálních komponentách a jsou součástí predikčního modelu. Model O2PLS je schopen pracovat se silnou korelovaností uvnitř matic  $X$  a  $Y$ , tzv. multikolinearitou a model stále funguje i v případě výskytu lineárních funkčních vztahů uvnitř těchto matic. Navíc model O2PLS nerozlišuje prediktory a predikované proměnné a je bez problémů schopen predikovat proměnné v matici  $Y$  na základě matice  $X$  i naopak (Meloun, Militký a Hill 2005, Stramová 2017).

V programu SIMCA P+ verze 12.0 byla analyzována data transformována směrem k symetrickému rozdělení a konstantnímu rozptylu (homoskedasticitě). Prvním krokem byly vyloučeny z analýzy parametry s více než 50 % chybějících údajů. Dalším krokem byla analýza s využitím vícenásobné regrese s redukcí dimenzionality (metoda obousměrných ortogonálních projekcí do latentní struktury, O2PLS). Byly vybrány matice závisle proměnných  $Y$  (diagnostická data) a nezávisle proměnných  $X$  (parametry životního prostředí). Následně proběhl algoritmus analýzy metodou O2PLS. Nejprve byla ověřena homogenita v matici  $X$  (prediktory, nezávisle proměnné) s využitím Hotellingovy  $T^2$  statistiky. Hotellingova  $T^2$  statistika je odvozena z Mahalanobisovy vzdálenosti a má F-rozdělení. Mahalanobisova vzdálenost, na rozdíl od vzdálenosti euklidovské, uvažuje také korelovanost mezi proměnnými. Odlehlé body detekované Hotellingovou  $T^2$  statistikou, jsou z dalších analýz vyřazeny. Na grafu 22 je zaznamenána Hotellingova  $T^2$  statistika, kde červená přímka označuje hladinu významnosti  $p < 0,01$ . Hladina významnosti z důvodu vysoké četnosti dat byla posunuta na hodnotu  $p < 0,002$ .



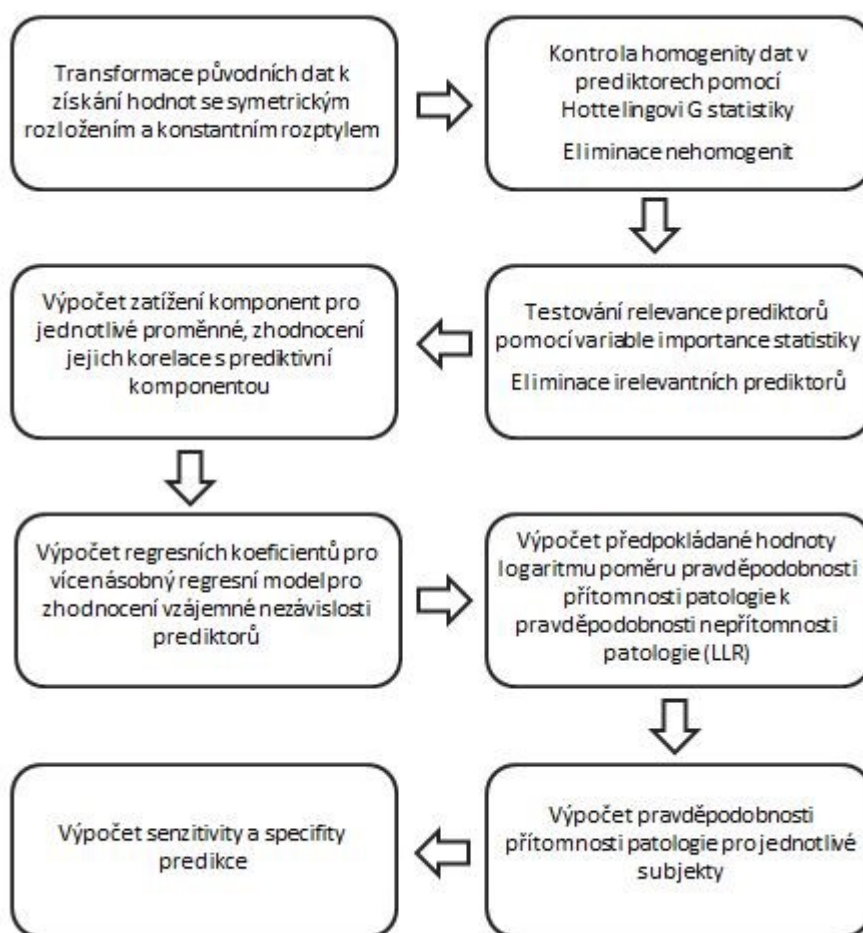
**Graf 22. Hottelingova  $T^2$  statistika**

Po vyřazení všech odlehlých bodů je opět spuštěna O2PLS a opět se zobrazí Hottelingovy  $T^2$  statistiky. Tento postup je opakovaně aplikován, dokud nejsou všechny body s  $p < 0,002$  vyjmuty z analýzy. Běžně nelze všechny nehomogenity detekovat napoprvé z důvodu tzv. maskování dalších nehomogenit nehomogenitami zvláště výraznými, teprve po vyjmutí extrémních nehomogenit se skryté nehomogenity projeví. Následně je použita statistika *VIP* (variable importance), aby byla ověřena relevance jednotlivých prediktorů. Do pokračující analýzy jsou pak zařazeny jen prediktory relevantní. Ty, které se významně ( $p < 0,05$ ) nepodílejí na konstrukci modelu, jsou z analýzy vyloučeny. Pokud by se do analýzy zahrnuly i nerelevantní proměnné, analýza by byla nucena „vysvětlit nevysvětlitelné“ a vnášel by se pouze další nevysvětlený rozptyl. V důsledku toho by se snižovala prediktivita modelu. *VIP* statistiky se testují ve všech prediktivních komponentách. Za nerelevantní prediktory se považují jen ty, které nenabudou významnosti u žádné z významných prediktivních komponent. Program SIMCA stanovuje významnost i počet prediktivních komponent v automatickém režimu.

V dalším kroku jsou opět testovány Hotellingovy  $T^2$  statistiky v upraveném modelu. Postup je ukončen v okamžiku, kdy nejsou nalezeny žádné experimentální body k odstranění a ani není nutné odebrat proměnné v matici prediktorů  $X$ . Poté je provedena záměna matice prediktorů  $X$  a matice predikovaných proměnných  $Y$ . Stejný postup, jaký byl u matice  $X$ , se aplikuje u matice závislé proměnných  $Y$ . Proces se ukončí v okamžiku, kdy v systému nejsou nerelevantní proměnné ani odlehlé body.

V dalším kroku se zamění matice  $X$  za  $Y$  a spouští se testování homogenity prediktorů a relevance proměnných. V okamžiku neexistence nerelevantních proměnných a nehomogenit v maticích  $X$  i  $Y$  je celý postup doveden ke konečnému výsledku. Následně se zobrazí rezidua (rozdíly mezi měřeními a predikovanými hodnotami) u závisle proměnných. Rezidua jednotlivých závislých proměnných byla zobrazena pomocí kvantil-kvantilového grafu a odlehlé body příslušející těmto reziduům byla z analýzy vyloučena. Algoritmus je úspěšně doveden do cíle v okamžiku, kdy neexistují nerelevantní proměnné v modelu O2PLS, nehomogenity v maticích  $X$  a  $Y$  a analýza reziduů nenachází nehomogenity v jednotlivých proměnných matice závisle proměnných.

Celý postup je názorně zachycen na Obr. 23.



**Obr. 23 Algoritmus pro analýzu výskytu nemoci s vnějšími faktory**

#### **4.4.1 O2PLS při nastavení závislé proměnné všech lékařsko diagnostických parametrů**

Prvním krokem byla volba matic závislé a nezávislé proměnných. Jako závislé proměnné  $Y$  byly zvoleny veškeré lékařsko-diagnostické parametry a jako nezávislé proměnné  $X$  pak ukazatele životního prostředí a pohlaví, věk, bmi a jodur. Byl aplikován celý postup, který je popsán výše, aby bylo možné analýzu ukončit a interpretovat výsledky.

Výsledkem tohoto postupu je vznik regresního modelu o jediné prediktivní komponentě. Tato komponenta vysvětluje 5,9 % variability modelu a 4,3% po křížové validaci. Ze závislé proměnných zde zbyla ze všech parametrů zdravotního stavu pacienta pouze jedna proměnná a tou byla přítomnost alergie (alergik\_LLRL). Vysokou míru vlivu ( $R$ ) mají proměnné PM10\_imise, BaP\_imise, SO2\_M4\_imise, Pb\_imise, Ni\_imise, Cd\_imise a to jak za místo bydliště tak zprůměrované s polohou školy a poté pr\_Be\_vod. U všech těchto parametrů životního prostředí je vysoká míra statistické jistoty ( $t$ -statistika). Celkově bylo identifikováno 51 parametrů, které významně ovlivňují model. Všechny tyto parametry s jejich míru vlivu a statistickou jistotou tohoto vlivu jsou uvedeny v Tab. 15.

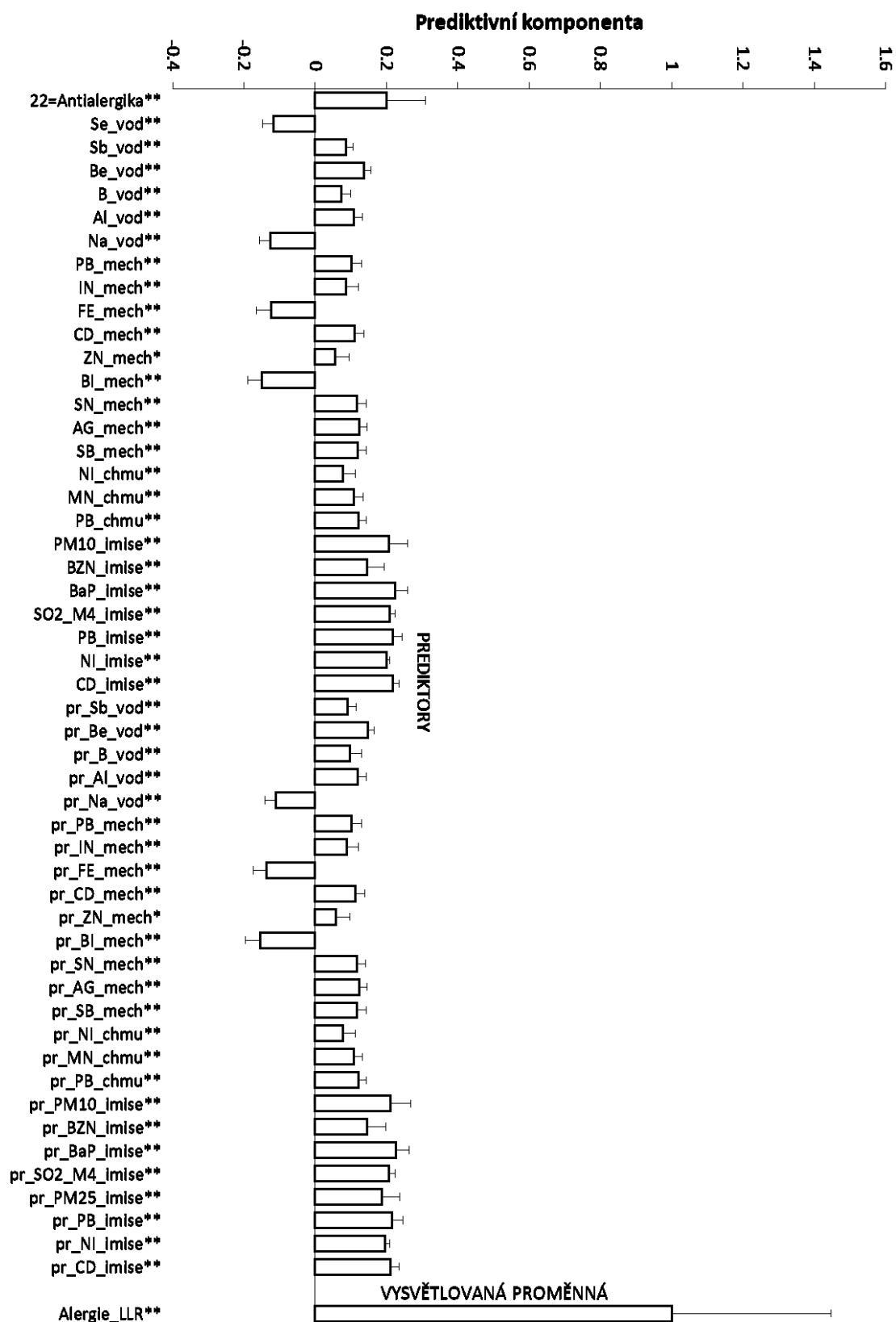
		Prediktivní komponenta		
Proměnné		Nakládací komponenta	t-statistika	R <sup>a</sup>
Relevantní prediktory (matice X)	22=Antialergika	0.199	3.40	0.475 **
	Se_vod	-0.117	-7.50	-0.455 **
	Sb_vod	0.086	8.33	0.362 **
	Be_vod	0.137	13.06	0.613 **
	B_vod	0.073	5.34	0.293 **
	Al_vod	0.109	8.89	0.439 **
	Na_vod	-0.126	-7.79	-0.501 **
	PB_mech	0.101	6.34	0.353 **
	IN_mech	0.086	4.73	0.296 **
	FE_mech	-0.123	-5.51	-0.447 **
	CD_mech	0.111	7.84	0.391 **
	ZN_mech	0.056	2.70	0.184 *
	BI_mech	-0.149	-6.95	-0.568 **
	SN_mech	0.117	8.57	0.412 **
	AG_mech	0.123	10.13	0.433 **
	SB_mech	0.118	9.05	0.418 **
	NI_chmu	0.079	4.29	0.269 **
	MN_chmu	0.109	8.26	0.383 **
	PB_chmu	0.122	10.91	0.435 **
	PM10_imise	0.206	7.23	0.799 **
	BZN_imise	0.145	5.58	0.568 **
	BaP_imise	0.225	12.68	0.857 **
	SO2_M4_imise	0.208	25.68	0.778 **
	PB_imise	0.217	15.83	0.824 **
	NI_imise	0.200	42.55	0.740 **
	CD_imise	0.217	22.59	0.813 **
	pr_Sb_vod	0.091	7.72	0.372 **
	pr_Be_vod	0.148	15.31	0.641 **
	pr_B_vod	0.097	5.78	0.376 **
	pr_Al_vod	0.119	9.42	0.474 **
	pr_Na_vod	-0.110	-6.94	-0.442 **
	pr_PB_mech	0.101	6.42	0.354 **
	pr_IN_mech	0.088	4.82	0.303 **
	pr_FE_mech	-0.136	-6.75	-0.493 **
	pr_CD_mech	0.112	8.01	0.394 **
	pr_ZN_mech	0.058	2.81	0.189 *
	pr_BI_mech	-0.155	-7.16	-0.590 **
	pr_SN_mech	0.116	8.58	0.412 **
	pr_AG_mech	0.123	10.59	0.437 **
	pr_SB_mech	0.118	8.96	0.417 **
	pr_NI_chmu	0.078	4.26	0.269 **
	pr_MN_chmu	0.108	8.04	0.379 **
	pr_PB_chmu	0.121	10.72	0.431 **
	pr_PM10_imise	0.210	6.89	0.813 **
	pr_BZN_imise	0.147	5.42	0.576 **
	pr_BaP_imise	0.227	11.53	0.862 **
	pr_SO2_M4_imise	0.206	21.60	0.770 **
	pr_PM25_imise	0.187	7.04	0.727 **
	pr_PB_imise	0.216	13.50	0.814 **
	pr_NI_imise	0.196	26.91	0.724 **
	pr_CD_imise	0.211	16.82	0.789 **
(matice Y)	Alergie_LLRL	1.000	4.24	0.243 **
Vysvětlená variabilita		5.9% (4.3% po křížové validaci)		
°R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01				

**Tab. 15. Vztah mezi ukazatelem alergie a vnějšími faktory hodnocený modelem O2PLS, prediktivní komponenta**



V grafu 23 jsou následně zobrazeny komponentní váhy první prediktivní komponenty, které jsou popsány pomocí regresních koeficientů s jejich 95 % intervaly spolehlivosti. Je zřejmé, že charakteristiky prostředí ovlivňují pozitivně i negativně pravděpodobnost výskytu alergie u pacientů. Zajímavé je, že ve výsledném modelu se neprojevuje vliv věku nebo pohlaví pacienta. Jediný významný vliv, který byl označen jako nezávisle proměnná a přitom nesouvisí s životním prostředím je užívání antialergetik. Nicméně užívání antialergik jednoznačně souvisí s přítomností alergie, neboť zdraví lidé antialergika neužívají.

Dalším krokem bylo vytvoření Tab. 16 a grafu vícenásobné regrese pro proměnnou alergik. V tabulce jsou dvěma hvězdičkami označené parametry, které mají vysokou statistickou jistotu ve vztahu k proměnné alergik. Toto označení je u parametrů obsah bismutu v mechu, benzenu v imisích, kadmia v imisích a to jak za místo bydliště, tak za průměrné hodnoty vzdálenosti s polohou školy a posledním významným parametrem je průměrná hodnota benzpyrenu v místě bydliště a školy. U těchto ukazatelů je také vysoká míra vlivu vyjádřená regresním koeficientem.

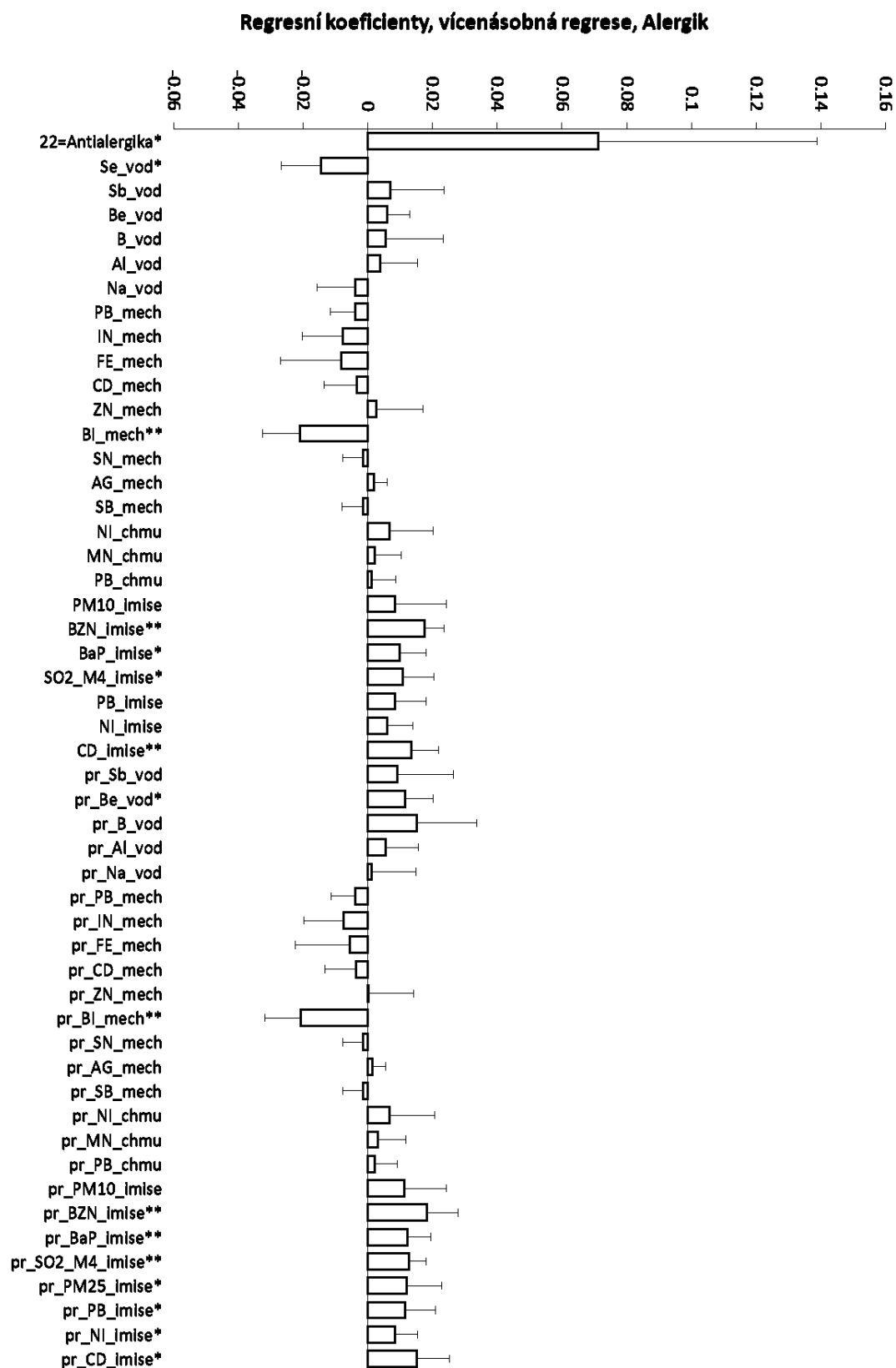


**Graf 23.** Komponentní váhy prediktivní komponenty vyjádřené jako regresní koeficienty s jejich 95 % intervaly spolehlivosti

		Prediktivní komponenta	
Proměnná		Regresní koeficient	t-statistiky
Relevantní prediktory (matice X)	22=Antialergika	0.071	1.99 *
	Se_vod	-0.014	-2.23 *
	Sb_vod	0.007	0.79
	Be_vod	0.006	1.60
	B_vod	0.005	0.58
	Al_vod	0.004	0.63
	Na_vod	-0.004	-0.60
	PB_mech	-0.004	-0.90
	IN_mech	-0.008	-1.15
	FE_mech	-0.008	-0.82
	CD_mech	-0.003	-0.64
	ZN_mech	0.003	0.33
	BI_mech	-0.021	-3.39 **
	SN_mech	-0.002	-0.45
	AG_mech	0.002	0.85
	SB_mech	-0.001	-0.41
	NI_chmu	0.007	0.98
	MN_chmu	0.002	0.49
	PB_chmu	0.001	0.32
	PM10_imise	0.008	1.01
	BZN_imise	0.017	5.39 **
	BaP_imise	0.010	2.27 *
	SO2_M4_imise	0.011	2.07 *
	PB_imise	0.008	1.64
	NI_imise	0.006	1.39
	CD_imise	0.013	3.01 **
	pr_Sb_vod	0.009	1.01
	pr_Be_vod	0.012	2.59 *
	pr_B_vod	0.015	1.54
	pr_Al_vod	0.006	1.06
	pr_Na_vod	0.001	0.17
	pr_PB_mech	-0.004	-0.96
	pr_IN_mech	-0.007	-1.15
	pr_FE_mech	-0.005	-0.61
	pr_CD_mech	-0.003	-0.68
	pr_ZN_mech	0.000	0.03
	pr_BI_mech	-0.021	-3.57 **
	pr_SN_mech	-0.001	-0.43
	pr_AG_mech	0.001	0.59
	pr_SB_mech	-0.001	-0.39
	pr_NI_chmu	0.007	0.93
	pr_MN_chmu	0.003	0.67
	pr_PB_chmu	0.002	0.60
	pr_PM10_imise	0.011	1.61
	pr_BZN_imise	0.018	3.57 **
pr_BaP_imise	0.012	3.24 **	
pr_SO2_M4_imise	0.013	4.54 **	
pr_PM25_imise	0.012	2.09 *	
pr_PB_imise	0.012	2.37 *	
pr_NI_imise	0.008	2.24 *	
pr_CD_imise	0.015	2.82 *	
Vysvětlená variabilita		5.9 % (4.3% po křížové validaci)	
* <i>p</i> <0.05, ** <i>p</i> <0.01			

Tab. 16. Vztah Alergie vs. Prediktory hodnocený vícenásobnou regresí

V grafu 24 je rovněž zobrazen vliv jednotlivých proměnných. Nejvyšší pozitivní korelace je zjištěna u antialergika, BZN\_imise, Bap\_imise, CO2\_imise, Cd\_imise, pr\_Be\_vod a průměrné hodnoty v imisích BZN, BaP, SO2, PM25, Pb, Ni a Cd. Znamená to tedy, že čím vyšší naměřená hodnota vyjmenovaných parametrů tím pravděpodobněji má pacient alergii. U proměnné alergik se nejedná o specifickou alergii na jednotlivé prvky, ale pouze o její přítomnost obecně. Lze však říci, že prvky s pozitivní korelací mohou být alergeny. Všechny jednotlivé parametry jsou podrobně popsány v předchozích kapitolách, kde je uveden vliv na zdraví člověka a u části je na základě údajů v literatuře pravděpodobné, že mohou být alergeny, což naše data rovněž ukazují. Z grafu je dále vidět, že negativní korelace existuje u hodnot koncentrace Se v pitné vodě a koncentrace Bi v mechu, která srovnatelně významná jako místo bydliště tak průměrná hodnota bydliště se školou. To znamená, že čím je menší hladina těchto parametrů tím větší pravděpodobnost přítomnosti alergie pacienti mají. Selen je pro člověka důležitý prvek, který hraje důležitou roli v lidském organismu při ochraně buněk a tkání před množstvím, které pronikne do lidského organismu. U bismutu není zjištěno, že by se jednalo o esenciální prvek pro lidský organismus. Jde o prvek s nízkou toxicitou. Jak bismut, tak selen jsou vzácnější prvky. Data o koncentraci selenu v pitné vodě jsou poměrně přesná a pokrývají velké množství území, avšak význam koncentrace selenu v regresním modelu je méně významný ve srovnání s koncentrací bismutu v mechu. Nicméně, tato data jsou určena spíše pro celorepublikový přehled a ne pro oblasti velikosti okresu. Bylo by tedy vhodné získat v budoucnu lepší datovou základnu koncentrace Bi v mechu i v imisích a pokusit se ověřit relevanci výsledků této analýzy.



**Graf 24. Výsledek vícenásobné regrese pro Alergie**

#### 4.4.2 OPLS při nastavení závislé proměnné VOLGUT

Při nastavení závisle proměnných  $Y$  u všech lékařsko-diagnostických parametrů byl výsledkem regresní model pro parametr alergie. Práce byla primárně zaměřena na parametry funkce štítné žlázy, proto byly jako prediktory postupně voleny pouze jednotlivé parametry funkce štítné žlázy. V první části byl jako závisle proměnná  $Y$  zvolen parametr VOLGUT a jako prediktory  $X$  byly zvoleny ukazatele životního prostředí a zbylé lékařsko-diagnostické parametry. Poté byl aplikován celý výše popsany postup.

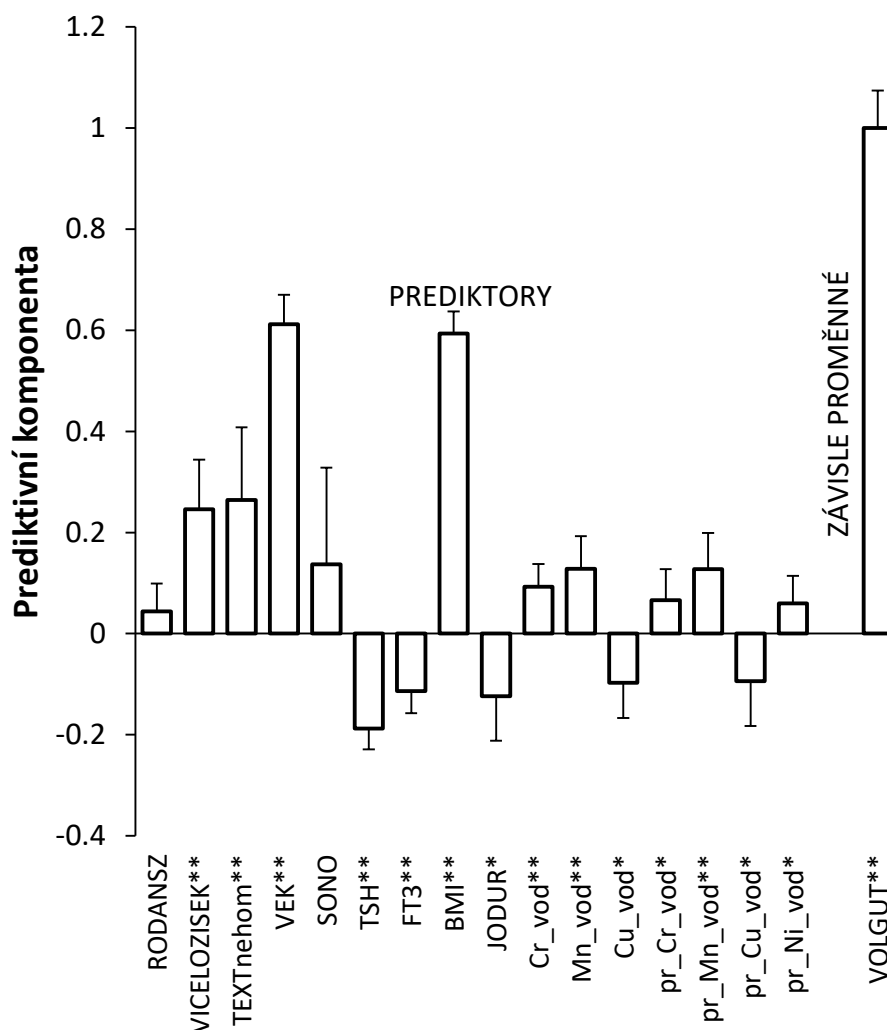
Při tomto postupu bylo vysvětleno ze 61,4 % (po křížové validaci z 60,3 %) variability parametru VOLGUT. Tato proměnná je významným diagnostickým ukazatelem. Na vysvětlení variability VOLGUT se podílejí další lékařsko-diagnostické proměnné i ukazatele životního prostředí a to hlavně v pitné vodě. Všechny prediktory jsou vyobrazeny v následující tabulce. Je zřejmé, že nejvíce je VOLGUT ovlivňován věkem. Veškeré výsledky jsou zaznamenány v Tab. 17.

Prediktivní komponenta				
	Proměnné	Nakládací komponenta	t-statistika	R <sup>a</sup>
Relevantní prediktory (matice X)	RODANSZ	0.044	1.51	0.055
	VICELÓZISEK	0.246	4.77	0.352 **
	TEXTnehom	0.264	3.49	0.375 **
	VEK	0.612	19.84	0.841 **
	SONO	0.137	1.36	0.187
	TSH	-0.187	-8.56	-0.263 **
	FT3	-0.113	-4.87	-0.155 **
	BMI	0.593	25.60	0.821 **
	JODUR	-0.124	-2.68	-0.174 *
	Cr_vod	0.093	3.90	0.083 **
	Mn_vod	0.128	3.73	0.130 **
	Cu_vod	-0.097	-2.66	-0.083 *
	pr_Cr_vod	0.066	2.03	0.048 *
	pr_Mn_vod	0.128	3.38	0.133 **
	pr_Cu_vod	-0.094	-2.00	-0.087 *
	pr_Ni_vod	0.060	2.09	0.052 *
(matice Y)	VOLGUT	1.000	25.50	0.783 **
Vysvětlená variabilita		61.4% (60.3% po křížové validaci)		
<sup>a</sup> R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01				

**Tab. 17. Vztah VOLGUT a vnější faktory hodnocený modelem OPLS**

Čím je pacient starší nebo má vyšší BMI tím má větší objem štítné žlázy. VOLGUT se podle grafu zvětšuje i s větším počtem ložisek, dále při nehomogenní textuře štítné žlázy. Z tabulky 17 a grafu 25 je zřejmé, že s VOLGUT pozitivně koreluje mangan a chrom. Přes pozitivní korelaci Cr v pitné vodě a VOLGUT, koncentrace Cr v pitné vodě nepřekračují žádnou z maximálních povolených hodnot. U Mn byla u několika obcí naměřena koncentrace,

která překračuje povolené mezní hodnoty. Vliv Cu v pitné vodě je opačný než u Mn a Cr. Cu negativně koreluje s VOLGUT.



**Graf 25.** Komponentní váhy prediktivní komponenty VOLGUT vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti

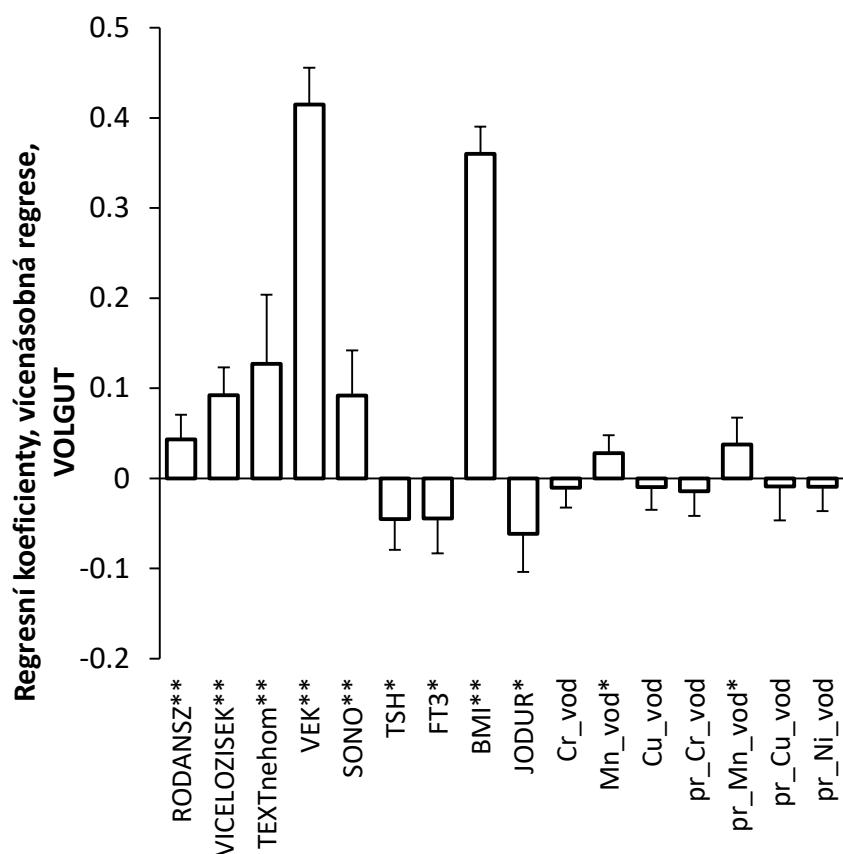
Pro VOLGUT také byla vytvořena tabulka (18) a graf (26) vícenásobné regrese. Z Tab. 18 je viditelné, že relevantními parametry pro VOLGUT jsou lékařsko-diagnostické ukazatele, z nichž nejvýznamnější jsou BMI a věk, které korelují pozitivně s objemem štítné žlázy. Tedy, že čím starší pacient a čím větší má hodnotu BMI tím větší bude pravděpodobně mít hodnotu VOLGUT. Když dojde ke zvětšení štítné žlázy, tak se jedná o závažný zdravotní problém. Zvětšená štítná žláza může způsobovat problémy při polykání a při dýchání, v určitých situacích může docházet k chrapotu, neboť zbytnělá žláza tlačí na nervy, které ovládají hlasivky.



Prediktivní komponenta			
Proměnné	Regresní koeficient	t-statistika	
Relevantní prediktory (matice X)	RODANSZ	0.043	3.01 **
	VICELZISEK	0.092	5.65 **
	TEXTnehom	0.127	3.13 **
	VEK	0.415	19.21 **
	SONO	0.092	3.46 **
	TSH	-0.045	-2.51 *
	FT3	-0.044	-2.16 *
	BMI	0.360	22.55 **
	JODUR	-0.061	-2.76 *
	Cr_vod	-0.010	-0.89
	Mn_vod	0.028	2.70 *
	Cu_vod	-0.010	-0.75
	pr_Cr_vod	-0.014	-1.00
	pr_Mn_vod	0.038	2.38 *
	pr_Cu_vod	-0.009	-0.44
pr_Ni_vod	-0.009	-0.65	
Vysvětlená variabilita		61.4 % (60.3% po křížové validaci)	
* <i>p</i> <0.05, ** <i>p</i> <0.01			

**Tab. 18. Vzath Volgut vs. Prediktory hodnocený vícenásobnou regresí**

Z výsledků vícenásobné regrese je patrné, že pro VOLGUT je nejvýznamnějším parametrem životního prostředí obsah manganu v pitné vodě. Vliv manganu na lidské tělo je popsán v části s názvem „Koncentrace MN v pitné vodě“. Důležité zde je, že lidské tělo v určité míře Mn potřebuje. Pro objem štítné žlázy se jedná o pozitivní korelaci s manganem, který je přítomný v pitné vodě. Z dalších parametrů pitné vody se zde projevil Cr, Cu a Ni. Ostatní ukazatele životního prostředí nemají dle tohoto modelu významný vliv na VOLGUT.



**Graf 26.** Výsledek vícenásobné regrese pro Volgut

#### 4.4.3 OPLS při nastavení závislé proměnné TSH

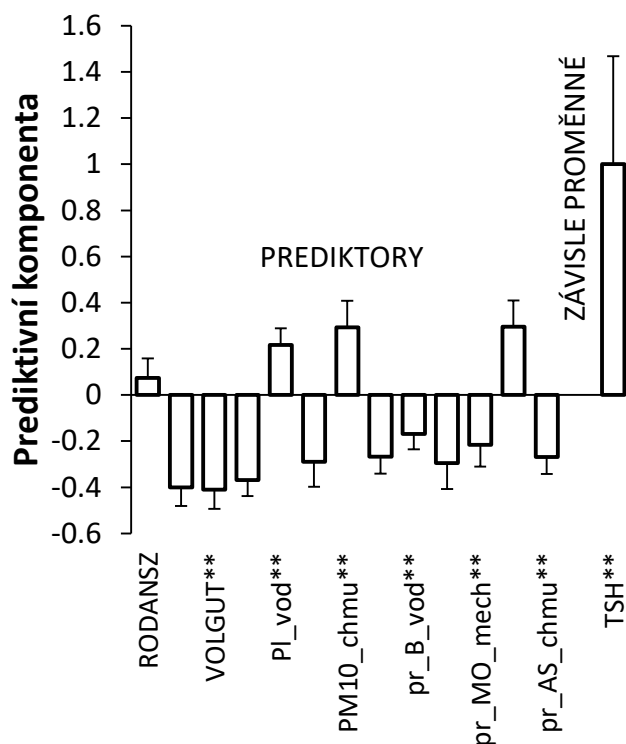
Dalším parametrem, na který byla aplikována metoda OPLS, byl hormon TSH. Stejně jako u VOLGUT byl zvolen stejný rejstřík závisle a nezávisle proměnných, pouze s tím rozdílem, že místo VOLGUT byl závisle proměnnou TSH.

Výsledkem byla prediktivní komponenta (Tab. 19), která byla tvořená 13 relevantními prediktory a závisle proměnnou. Prediktory vysvětlovaly 10,3 % variability TSH a po křížové validaci pak 8,7 %. Oproti VOLGUT je u TSH podíl vysvětlené variability nižší, ale korelace mezi maticemi  $X$  i  $Y$  je stále vysoce významná. U TSH se méně projevil vliv lékařsko-diagnostických parametrů a více vliv ukazatelů životního prostředí. Nejvýznamnějšími z lékařsko-diagnostických ukazatelů byly opět parametry BMI, věk a VOLGUT. Z parametrů životního prostředí se zde projevil vliv pesticidních látek, Mn a B v pitné vodě, PM10 a As v ovzduší a také vliv Mo v mechu.

Prediktivní komponenta				
	Proměnné	Nakládací komponenta	t-statistika	R <sup>2</sup>
Relevantní prediktory (matice X)	RODANSZ	0.073	1.62	0.121
	VEK	-0.401	-9.46	-0.660 **
	VOLGUT	-0.411	-9.48	-0.677 **
	BMI	-0.368	-10.05	-0.606 **
	PI_vod	0.217	5.77	0.347 **
	Mn_vod	-0.289	-5.06	-0.460 **
	PM10_chmu	0.293	4.84	0.482 **
	AS_chmu	-0.268	-6.91	-0.441 **
	pr_B_vod	-0.169	-4.73	-0.272 **
	pr_Mn_vod	-0.295	-4.98	-0.470 **
	pr_MO_mech	-0.217	-4.39	-0.357 **
	pr_PM10_chmu	0.296	4.93	0.487 **
	pr_AS_chmu	-0.268	-6.85	-0.442 **
(matice Y)	TSH	1.000	4.05	0.320 **
Vysvětlená variabilita		10.3% (8.7% po křížové validaci)		
ªR...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01				

**Tab. 19. Vztah TSH a vnější faktory hodnocený modelem OPLS**

Graficky znázorněná komponenta je na grafu 27 a znázorňuje pozitivní či negativní korelaci jednotlivých ukazatelů. Největší pozitivní korelaci s TSH měly PM10 v ovzduší, dále pesticidní látky ve vodě a slabší korelaci s TSH pak vykazovala léčba příbuzných se štítnou žlázou. Naopak negativně korelovaly s TSH vliv všechny zbylé relevantní parametry.



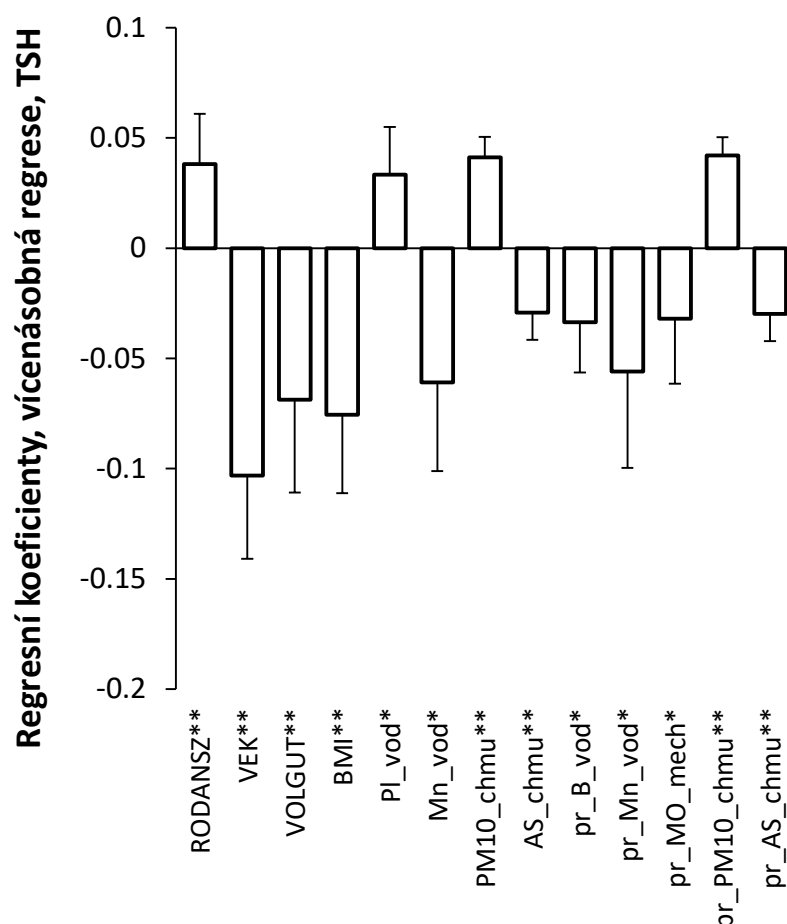
**Graf 27. Komponentní váhy prediktivní komponenty TSH vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti**

Následně byla použita vícenásobná regrese, na jejímž základě byla vytvořena Tab. 20. Z té jsou patrné jednotlivé regresní koeficienty výsledných parametrů. Z tabulky je patrná velká statistická jistota pro vliv koncentrací PM10 a As v ovzduší a to jak za místo bydliště, tak za zprůměrované hodnoty s místem školy. Je proto pravděpodobné, že tyto ukazatele mohou souviset s hladinami TSH. Dalšími významnými parametry jsou věk, BMI a rodinná anamnéza.

		Prediktivní komponenta	
	Proměnné	Regresní koeficient	t-statistika
Relevantní prediktory (matice X)	RODANSZ	0.038	3.17 **
	VEK	-0.103	-5.17 **
	VOLGUT	-0.069	-3.10 **
	BMI	-0.076	-4.02 **
	Pl_vod	0.033	2.91 *
	Mn_vod	-0.061	-2.86 *
	PM10_chmu	0.041	8.38 **
	AS_chmu	-0.029	-4.51 **
	pr_B_vod	-0.034	-2.80 *
	pr_Mn_vod	-0.056	-2.43 *
	pr_MO_mech	-0.032	-2.04 *
	pr_PM10_chmu	0.042	9.67 **
pr_AS_chmu	-0.030	-4.60 **	
Vysvětlená variabilita		10.3 % (8.7% po křížové validaci)	
* <i>p</i> <0.05, ** <i>p</i> <0.01			

**Tab. 20. Vztah TSH vs. Prediktory hodnocený vícenásobnou regresí**

Z grafu 28 vícenásobné regrese je zřejmá pozitivní korelace s TSH u rodinné anamnézy. Pokud měl někdo z příbuzných pacienta onemocnění funkce štítné žlázy tak je pravděpodobnější, že pacient bude mít také problém se štítnou žlázou a tedy nadprodukcí TSH. Naopak velmi výrazné negativní korelace jsou patrné u věku, VOLGUT, BMI, které v regresním modelu a převyšují vliv ukazatelů životního prostředí. Z ukazatelů životního prostředí je patrná negativní korelace arsenu v ovzduší s TSH. Významný vliv mají také Mn a B v pitné vodě.



**Graf 28.** Výsledek vícenásobné regrese pro TSH

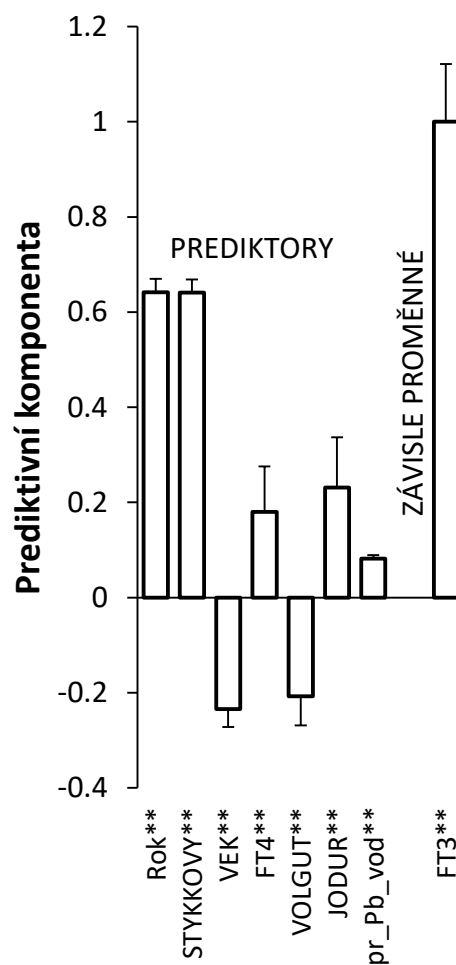
#### 4.4.4 OPLS při nastavení závislé proměnné FT3

Metoda OPLS byl dále použita pro predikci FT3. FT3 byla zvolena jako jediná závisle proměnná, zatímco všechny ostatní parametry byly zvoleny jako prediktory. Výsledkem je prediktivní komponenta, významně korelující se sedmi relevantními (Tab. 21). Z těchto prediktorů je pouze jeden ukazatelem životního prostředí a to průměrná hodnota olova v pitné vodě za místo bydliště a školu pacienta. Ostatní parametry jsou tedy lékařsko-diagnostické. U všech prediktorů je vysoká míra vlivu ( $R$ ) a to na hladině pravděpodobnosti  $p < 0,01$ . Je zde také viditelná velmi vysoká míra statistické jistoty. Celková vysvětlená variabilita FT3 je 38,3 % a po křížové validaci to je 37,6 %, což jsou velmi dobré výsledky, na kterých se podílí všechny prediktory velkou mírou.

Prediktivní komponenta					
	Proměnné	Nakládací komponenta	t-statistika	R <sup>a</sup>	
Relevantní prediktory (matice X)	Rok	0.642	43.66	0.929	**
	STYKKOVY	0.641	43.31	0.927	**
	VEK	-0.234	-11.71	-0.340	**
	FT4	0.180	3.57	0.263	**
	VOLGUT	-0.208	-6.46	-0.300	**
	JODUR	0.231	4.14	0.327	**
	pr_Pb_vod	0.082	20.91	0.118	**
(matice Y)	FT3	1.000	15.65	0.619	**
Vysvětlená variabilita		38.3% (37.6% po křížové validaci)			
<sup>a</sup> R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01					

**Tab. 21. Vztah FT3 a vnější faktory hodnocený modelem OPLS**

Z grafu 29 komponentních vah je vidět, že s pozdějším rokem vyšetření a s větším stykem s kovy roste hladina FT3 v těle pacienta. FT3 také pozitivně koreluje se s FT4, jodurii a koncentrací Pb v pitné vodě. Naopak FT3 negativně koreluje s věkem a objemem štítné žlázy. Čím vyšší jsou věk a objem štítné žlázy tím nižší bude pravděpodobně hladina FT3.



**Graf 29. Komponentní váhy prediktivní komponenty FT3 vyjádřené jako regresní koeficienty s jejich 95 % intervaly spolehlivosti**

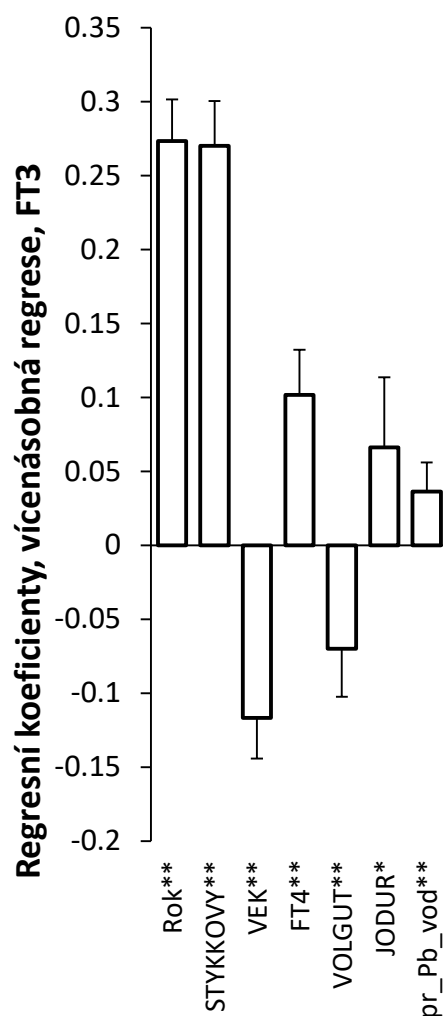
Následně byla vytvořena tabulka (22) a graf (30) pro model vícenásobné regrese. Z Tab. 22 je zřetelná významnost prediktorů FT3. Míra statistické jistoty klesla o parametru jodurie, ale přesto je stále významná ( $p < 0,05$ ). Stále je zde velmi významný rok vyšetření pacienta, který kladně koreluje s hladinou hormonu FT3. Tento výsledek je překvapivý a pro vysvětlení by bylo nutné porovnat obě skupiny vyšetřených (rok 2000 a rok 2005) mezi sebou. Pozitivní korelace mezi FT4 a FT3 je očekávaná neboť FT4 je prekurzorem FT3.

		Prediktivní komponenta	
	Proměnné	Regresní koeficient	t-statistika
Relevantní prediktory (matice X)	Rok	0.273	18.46 **
	STYKKOVY	0.270	16.80 **
	VEK	-0.117	-8.03 **
	FT4	0.102	6.34 **
	VOLGUT	-0.070	-4.08 **
	JODUR	0.066	2.64 *
	pr_Pb_vod	0.036	3.51 **
	Vysvětlená variabilita		38.3 % (37.6% po křížové validaci)
* <i>p</i> <0.05, ** <i>p</i> <0.01			

**Tab. 22. Vztah FT3 vs. Prediktory hodnocený vícenásobnou regresí**

Tabulka i graf ukazují poněkud silnější pozitivní korelace relevantních prediktorů, ale i o něco slabší, avšak stále významné negativní korelace prediktorů. Nejvyšší regresní koeficienty byly nalezeny u roku vyšetření a styku s kovy. U koncentrace olova v pitné vodě, která je zprůměrovaná za místo bydliště a školy, je vidět, že jeho vliv má poměrně vysokou míru statistické jistoty, ale jeho regresní koeficient, tedy míra vlivu, je nejnižší ze všech prediktorů. U olova je známá jeho toxicita a negativní působení na organismus člověka, a to zejména při vysokých koncentracích. V okrese Příbram byla naměřena zvýšená koncentrace olova v mechu a v ovzduší pak byla ve městě Příbram naměřena nejvyšší koncentrace v celé ČR. V pitné vodě olovo nepřesahovalo povolené mezní hodnoty, nicméně, pitná voda je měřena přímo u jejího zdroje a neměří se v domácnostech z kohoutku. Zde může nastávat problém, protože hlavním zdrojem olova v pitné vodě z kohoutku mohou být právě olověné trubky, kterými voda protéká (Kožíšek 2009). Z toho vyplývá, že v okrese Příbram jsou obyvatelé vystaveni relativně vyšší koncentraci olova, která může nějakým způsobem souviset s vyššími hladinami FT3.





**Graf 30. Výsledek vícenásobné regrese pro FT3**

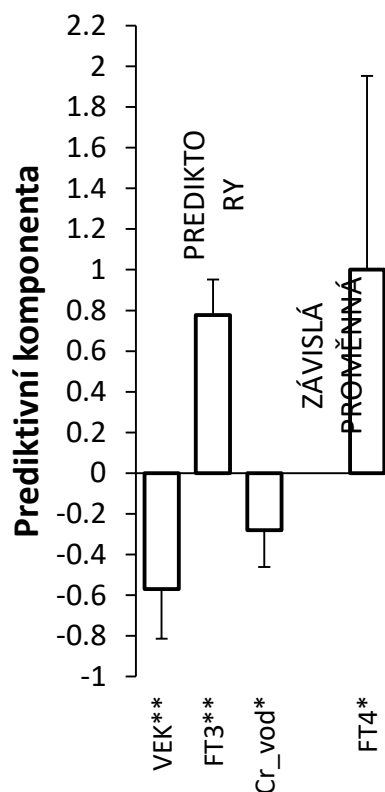
#### 4.4.5 OPLS při nastavení závislé proměnné FT4

Z analýzy OPLS při volbě FT4 jako závislé proměnné byla výsledkem jedna prediktivní komponenta zahrnující tři relevantní prediktory (Tab. 23). Vysvětlená variabilita u této komponenty je 5,6 % a po křížové validaci 4,5 %. Což je ve srovnání s předchozími parametry funkce štítné žlázy nižší hodnota, která má však rovněž důležitou výpovědní hodnotu.

<u>Prediktivní komponenta</u>				
	Variable	Nakládací komponenta	t-statistika	R <sup>a</sup>
Relevantní prediktory (matice X)	VEK	-0.569	-4.40	-0.621 **
	FT3	0.777	8.38	0.847 **
	Cr_vod	-0.280	-2.92	-0.302 *
(matice Y)	FT4	1.000	1.99	0.237 *
Vysvětlená variabilita		5.6% (4.5% po křížové validaci)		
<sup>a</sup> R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01				

**Tab. 23. Vztah FT4 a vnější faktory hodnocený modelem OPLS**

Z výsledků je patrná očekávaná pozitivní korelace s hladinou hormonu FT3. Tato korelace byla patrná i u výsledků metody OPLS pro závisle proměnnou FT3 s FT4 jako jedním z prediktorů. Jedná se tedy o potvrzení předchozích výsledků. Pro FT4 je FT3 relevantním prediktorem, který má největší míru vlivu i míru statistické jistoty a zároveň je to jediný parametr s pozitivní korelací. Negativní korelace jsou zřejmé z grafu 31 u zbylých parametrů tedy u věku a koncentrace chromu v pitné vodě.



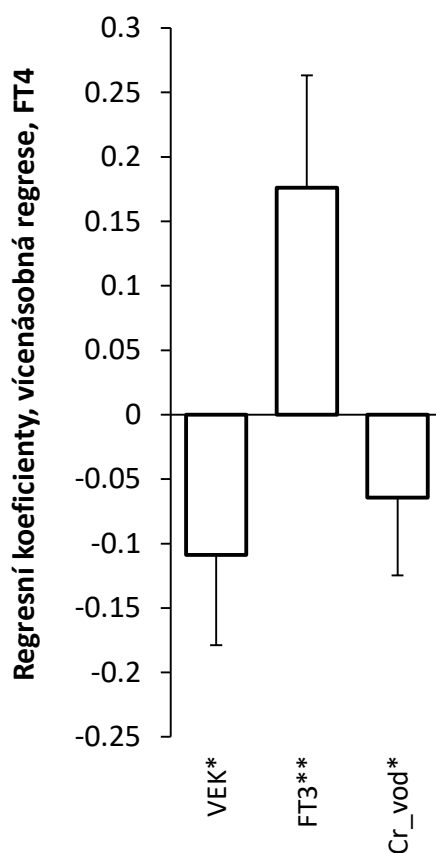
**Graf 31.** Komponentní váhy prediktivní komponenty FT4 vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti

Následně byla provedena regresní analýza daných parametrů. Z grafu (32) i tabulky (24) vyplývá nejvyšší hodnota regresního koeficientu i statistická jistota byla nalezena u prediktoru FT3 ( $p < 0,01$ ). Nižší statistická jistota byla zjištěna u věku a chromu ve vodě je dosaženo hladiny pravděpodobnosti ( $p < 0,05$ ). Uvedené relevantní prediktory vysvětlují 5,6 % variability FT4.

		Prediktivní komponenta	
		Regresní koeficient	t-statistika
Proměnné			
Relevantní prediktory (matice X)	VEK	-0.109	-2.95 *
	FT3	0.176	3.82 **
	Cr_vod	-0.064	-2.03 *
Vysvětlená variabilita		5.6 % (4.5% po křížové validaci)	
* $p<0.05$ , ** $p<0.01$			

**Tab. 24.** Vztah FT4 vs. Prediktory hodnocený vícenásobnou regresí

Z grafu 32 je vidět velikost regresních koeficientů pro jednotlivé parametry. Podobně jako FT3 také FT4 negativně koreluje s věkem. Ostatní parametry, které byly výsledkem OPLS analýzy pro FT3 (rok vyšetření, styk s kovy, jodurie, VOLGUT, Pb v pitné vodě) nebyly pro predikci FT4 relevantní. Z ukazatelů životního prostředí byla u FT4 významná negativní korelace průměrné koncentrace chromu v pitné vodě za místo bydliště pacienta. Tento ukazatel ovlivňuje FT4 nejméně avšak statisticky významně. Cr je významný esenciální prvek pro člověka, ale při vyšších koncentracích je alergizující a karcinogenní. V okrese Příbram byla naměřena na dvou místech zvýšená koncentrace Cr v mechu, ale v pitné vodě byly naměřeny koncentrace, které nepřesahovaly přípustné mezní hodnoty.



**Graf 32.** Výsledek vícenásobné regrese pro FT4

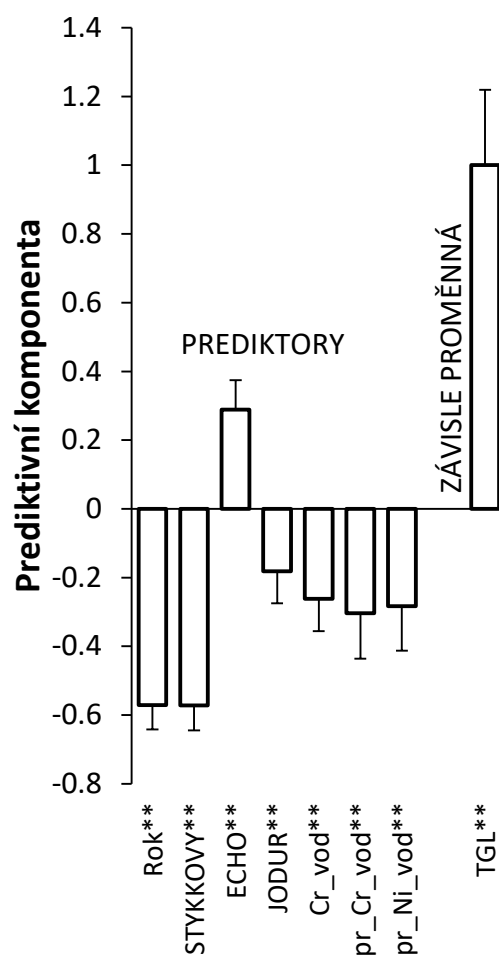
#### 4.4.6 OPLS při nastavení závislé proměnné TGL

V další řadě se přistoupilo k aplikaci metody OPLS na thyreoglobulin (TGL). Jedná se o glykoprotein, který obsahuje jód a za normálních podmínek je syntetizován ve folikulárních buňkách štítné žlázy (Pikner 2005). Výsledkem analýzy je komponenta zahrnující 7 relevantních prediktorů (Tab. 25). Tyto prediktory vysvětlují variabilitu TGL z 20,2 % a po křížové validaci z 19,2 %. Což je již středně silný vztah. Mezi těmito prediktory byly dva ukazatele životního prostředí.

Prediktivní komponenta					
Proměnné		Nakládací komponenta	t-statistika	R <sup>o</sup>	
Relevantní prediktory (matice X)	Rok	-0.571	-15.19	-0.885	**
	STYKKOVY	-0.572	-14.96	-0.886	**
	ECHO	0.289	6.36	0.447	**
	JODUR	-0.182	-3.71	-0.282	**
	Cr_vod	-0.262	-5.32	-0.414	**
	pr_Cr_vod	-0.304	-4.36	-0.479	**
	pr_Ni_vod	-0.283	-4.14	-0.447	**
(matice Y)	TGL	1.000	8.62	0.449	**
Vysvětlená variabilita		20.2% (19.2% po křížové validaci)			
°R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01					

**Tab. 25. Vztah TGL a vnější faktory hodnocený modelem OPLS**

Z tabulky (25) a grafu (33) je patrná pozitivní korelace pouze u parametru ECHO. U tohoto parametru byla zjištěná velká míra vlivu i statistické jistoty. Nejvyšší míra vlivu ale byla zjištěna u negativně korelujících parametrů rok vyšetření a styk s kovy. Tyto dva parametry hrají velkou roli i u analýzy FT3, kde mají rovněž velkou míru vlivu i statistické jistoty. Dále negativní korelace zřejmá také u jodurie a z ukazatelů životního prostředí pak u chromu v pitné vodě a to jak samostatně za místo bydliště tak průměrně za polohu bydliště a polohu školy, a v poslední řadě průměrná hodnota niklu v pitné vodě za polohu bydliště a školy.



**Graf 33.** Komponentní váhy prediktivní komponenty TGL vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti

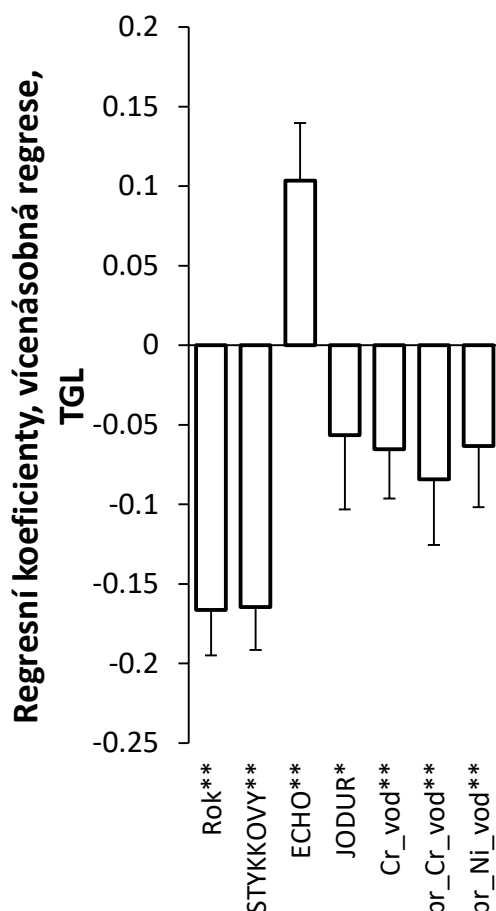
Při analýze vícenásobnou regresí opět vykazovaly nejvyšší regresní koeficienty rok vyšetření a styk s kovy a opět vysoce významný parametr ECHO (Tab. 26). Ukazatele životního prostředí mají nižší regresní koeficienty, nicméně jsou opět vysoce významné ( $p < 0,01$ ).

		Prediktivní komponenta	
	Proměnné	Regresní koeficient	t-statistika
Relevantní prediktory (matice X)	Rok	-0.166	-11.03 **
	STYKKOVY	-0.165	-11.61 **
	ECHO	0.103	5.42 **
	JODUR	-0.057	-2.29 *
	Cr_vod	-0.065	-4.01 **
	pr_Cr_vod	-0.084	-3.86 **
	pr_Ni_vod	-0.063	-3.12 **
Vysvětlená variabilita		20.2 % (19.2% po křížové validaci)	
* <i>p</i> <0.05, ** <i>p</i> <0.01			

**Tab. 26.** Vztah TGL vs. Prediktory hodnocený vícenásobnou regresí

Z grafu 34 je rovněž patrná negativní korelace ukazatelů životního prostředí, tedy čím vyšší jsou koncentrace Cr a Ni v pitné vodě tím nižší produkce TGL. U chromu již bylo řečeno, že se

jedná o esenciální prvek tj. prvek důležitý prvek pro lidský organismus, ale zároveň bylo řečeno, že jsou podstatné dávky, v jakých člověk chrom přijímá. U niklu se zatím neprokázala jeho nutnost pro lidský organismus, naopak jsou známe problémy s ním spojené a to právě při požití v pitné vodě. V okrese Příbram byly naměřeny hodnoty niklu v pitné vodě maximálně do 9,3 µg/l, což je přípustná koncentrace, která nepřekračuje povolený maximální limit (20 µg/l).



**Graf 34.** Výsledek vícenásobné regrese pro TGL

#### 4.4.7 OPLS při nastavení závislé proměnné RAS

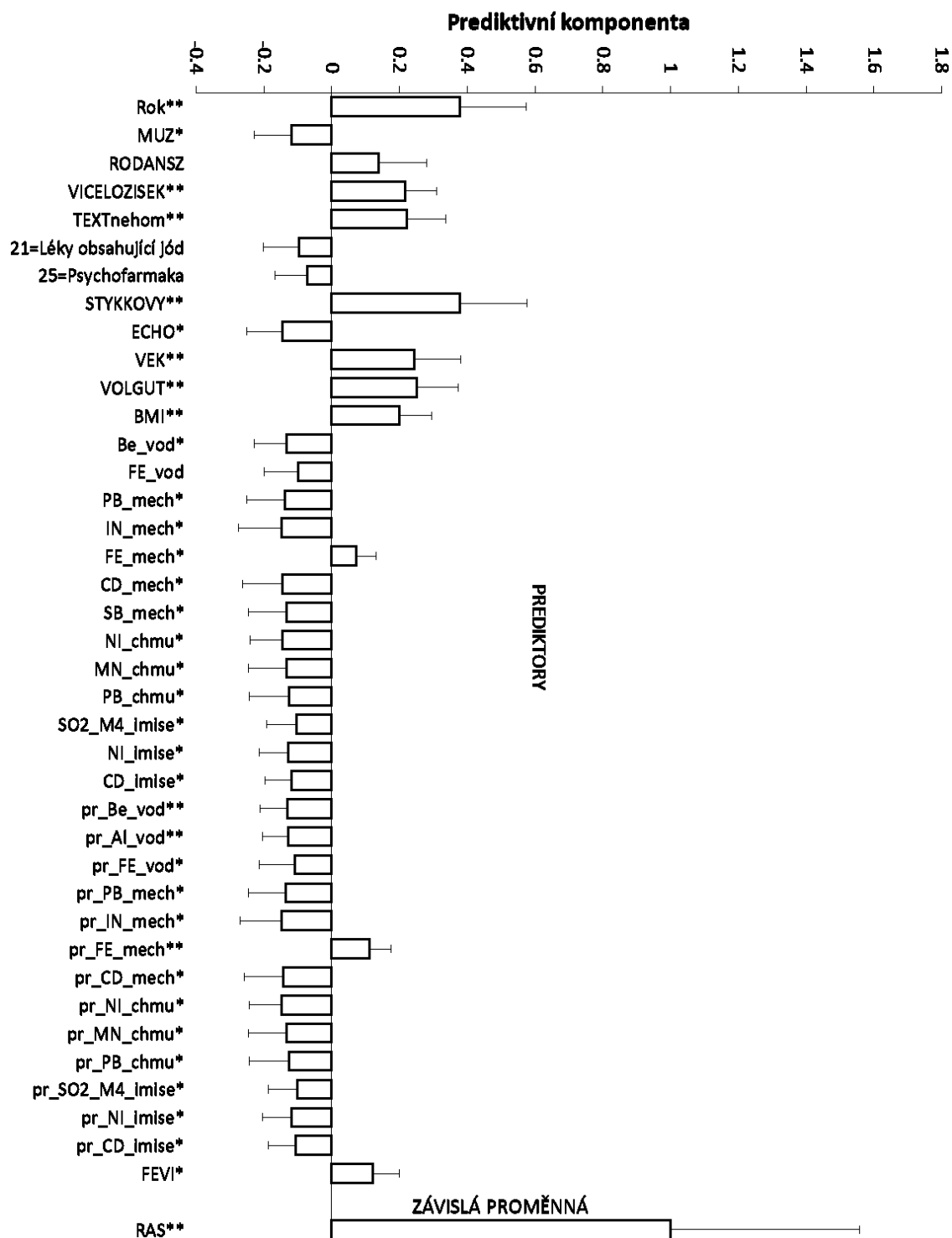
Posledním analyzovaným parametrem byla reflexe Achillovy šlachy (RAS). Jedná se o jeden z možných parametrů jak vyšetřit určité formy onemocnění funkce štítné žlázy a na základě délky trvání reflexu Achillovy šlachy také kontrolovat léčbu (Zamrazil 2002). Prediktivní komponenta modelu OPLS zahrnovala 39 relevantních prediktorů (Tab. 27). Ne všechny však mají stejně vysokou míru vlivu a statistické jistoty. Mezi nejvýznamnější prediktory patří: rok vyšetření, počet ložisek, textura štítné žlázy, styk s kovy, věk, VOLGUT, BMI, Be a Al v pitné vodě a Fe v mechu. Celkově se OPLS modelu podařilo vysvětlit variabilitu RAS z 13,9 % a po křížové validaci z 10,3 %.

Prediktivní komponenta				
	Proměnné	Nakládací komponenta	t-statistika	R <sup>2</sup>
Relevantní prediktory (matice X)	Rok	0.378	3.67	0.698 **
	MUZ	-0.118	-2.02	-0.215 *
	RODANSZ	0.138	1.83	0.255
	VICELOZISEK	0.216	4.33	0.399 **
	TEXTnehom	0.221	3.66	0.409 **
	21=Léky obsahující	-0.096	-1.70	-0.181
	25=Psychofarmaka	-0.071	-1.40	-0.132
	STYKKOVY	0.379	3.66	0.700 **
	ECHO	-0.145	-2.57	-0.267 *
	VEK	0.244	3.38	0.452 **
	VOLGUT	0.250	3.88	0.461 **
	BMI	0.199	3.92	0.367 **
	Be_vod	-0.133	-2.64	-0.291 *
	FE_vod	-0.098	-1.86	-0.168
	PB_mech	-0.139	-2.38	-0.255 *
	IN_mech	-0.148	-2.23	-0.272 *
	FE_mech	0.072	2.26	0.132 *
	CD_mech	-0.145	-2.32	-0.266 *
	SB_mech	-0.134	-2.27	-0.246 *
	NI_chmu	-0.144	-2.84	-0.264 *
	MN_chmu	-0.132	-2.20	-0.242 *
	PB_chmu	-0.126	-2.05	-0.231 *
	SO2_M4_imise	-0.103	-2.23	-0.189 *
	NI_imise	-0.128	-2.83	-0.233 *
	CD_imise	-0.119	-2.93	-0.218 *
	pr_Be_vod	-0.130	-3.06	-0.283 **
	pr_AI_vod	-0.127	-3.10	-0.269 **
	pr_FE_vod	-0.108	-1.91	-0.191 *
	pr_PB_mech	-0.136	-2.35	-0.249 *
	pr_IN_mech	-0.147	-2.27	-0.270 *
	pr_FE_mech	0.112	3.33	0.206 **
	pr_CD_mech	-0.143	-2.34	-0.262 *
	pr_NI_chmu	-0.147	-2.93	-0.269 *
	pr_MN_chmu	-0.133	-2.22	-0.243 *
	pr_PB_chmu	-0.125	-2.01	-0.228 *
	pr_SO2_M4_imise	-0.101	-2.24	-0.184 *
	pr_NI_imise	-0.119	-2.64	-0.217 *
	pr_CD_imise	-0.107	-2.53	-0.195 *
	FEVI	0.122	2.96	0.226 *
	(matice Y) RAS	1.000	3.40	0.373 **
Vysvětlená variabilita		13.9% (10.3% po křížové validaci)		
° R...zatížení komponent vyjádřené jako korelační koeficienty s prediktivní složkou, *p<0.05, **p<0.01				

**Tab. 27. Vztah RAS a vnější faktory hodnocený modelem OPLS**

Z tabulky (27) a grafu (35) je patrná významnost řady parametrů, nicméně převládají zde spíše negativní korelace s RAS. Parametry, které dosahují vysokou míru vlivu i statistické jistoty, mají pozitivní korelaci s RAS a parametry, které spíše celý model doplňují, s RAS většinou korelují negativně. Výjimkou je právě průměrná koncentrace Be a AI v pitné vodě za polohu bydliště a školy, což jsou významné parametry, které se velkou mírou podílejí na vysvětlené variabilitě, a mají negativní korelaci.





**Graf 35.** Komponentní váhy prediktivní komponenty RAS vyjádření jako regresní koeficienty s jejich 95 % intervaly spolehlivosti

Při regresní analýze byly vypočítány regresní koeficienty pro relevantní parametry. Na základě těchto regresních koeficientů je vidět, že významné parametry ubyly a významnosti dosáhlo pouze 14 z celkových 39 (Tab. 28). Výrazně převládají lékařsko-diagnostické parametry

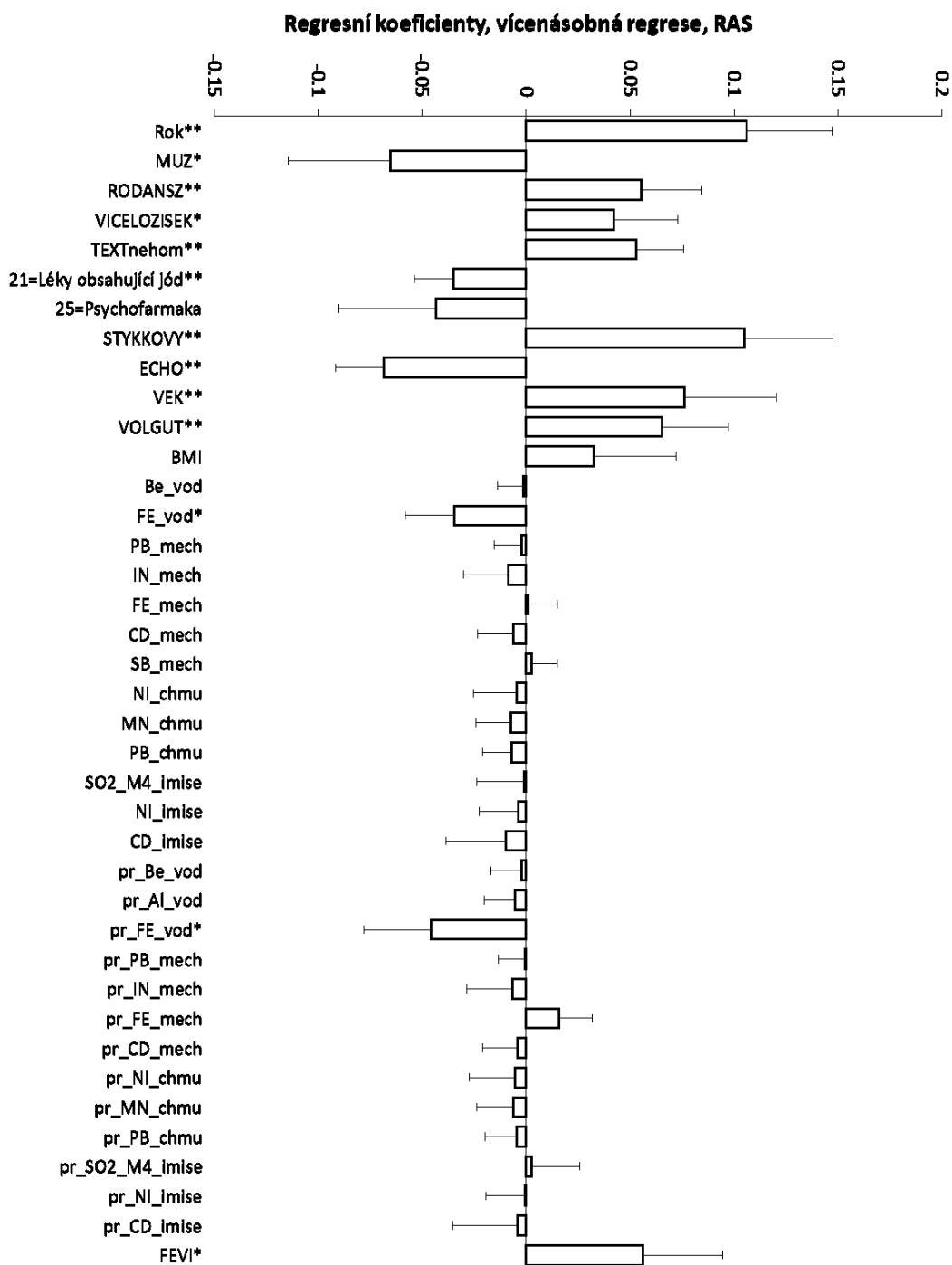
člověka a to nejvíce ECHO, textura a rok vyšetření. Rok vyšetření už byl diskutován v souvislosti s analýzou FT3. ECHO koreluje podle očekávání s RAS negativně, zatímco textura pozitivně. Když je textura štítné žlázy nehomogenní, je reflexe Achillovy šlachy delší.

Proměnné	Prediktivní komponenta	
	Regresní koeficient	t-statistika
Rok	0.106	4.88 **
MUZ	-0.065	-2.50 *
RODANSZ	0.055	3.62 **
VICELIZSEK	0.042	2.61 *
TEXTnehom	0.053	4.40 **
21=Léky obsahující	-0.035	-3.54 **
25=Psychofarmaka	-0.043	-1.75
STYKKOVY	0.105	4.66 **
ECHO	-0.069	-5.65 **
VEK	0.076	3.24 **
VOLGUT	0.065	3.87 **
BMI	0.033	1.56
Be_vod	-0.001	-0.22
FE_vod	-0.035	-2.77 *
PB_mech	-0.002	-0.32
IN_mech	-0.008	-0.74
FE_mech	0.001	0.17
CD_mech	-0.006	-0.66
SB_mech	0.003	0.39
NI_chmu	-0.005	-0.42
MN_chmu	-0.007	-0.82
PB_chmu	-0.007	-0.92
SO2_M4_imise	-0.001	-0.08
NI_imise	-0.004	-0.37
CD_imise	-0.010	-0.63
pr_Be_vod	-0.002	-0.28
pr_AI_vod	-0.005	-0.68
pr_FE_vod	-0.046	-2.65 *
pr_PB_mech	0.000	-0.01
pr_IN_mech	-0.007	-0.58
pr_FE_mech	0.016	1.88
pr_CD_mech	-0.004	-0.46
pr_NI_chmu	-0.005	-0.45
pr_MN_chmu	-0.006	-0.64
pr_PB_chmu	-0.005	-0.60
pr_SO2_M4_imise	0.002	0.20
pr_NI_imise	0.000	-0.02
pr_CD_imise	-0.004	-0.24
FEVI	0.056	2.78 *
25=Psychofarmaka	0.056	2.78 *
<b>Vysvětlená variabilita</b>		13.9 % (10.3% po křížové validaci)
* $p < 0.05$ , ** $p < 0.01$		

**Tab. 28. Vztah RAS vs. Prediktory hodnocený vícenásobnou regresí**

Z ukazatelů životního prostředí je pro model významná pouze průměrná koncentrace železa v pitné vodě a to samostatně jak za místo bydliště tak průměrně za polohu bydliště a školy (Graf 36). Železo je pro člověka esenciálním prvkem, který je součástí mnoha biologicky funkčních metaloproteinů – enzymů a cytochromů. Zároveň se ale ve vyšších koncentracích

může jednat o toxickou látku, která poškozuje játra a ledviny (Sucharová a kol. 2001). Z těchto výsledků se jedná o negativní korelaci Fe v pitné vodě s reflexem Achillovy šlachy. Tedy čím vyšší koncentrace Fe v pitné vodě tím nižší reflexe Achillovy šlachy. Koncentrace železa v pitné vodě V okrese Příbram nepřekračovala maximální mezní hodnoty, naopak v mechu byly identifikovány místa se zvýšenou koncentrací a okres Příbram se zařadil mezi místa s největší koncentrací Fe v mechu v celé ČR.



**Graf 36. Výsledek vícenásobné regrese pro RAS**

## 5 Diskuse výsledků

V této diplomové práci byl navržen metodický postup zpracování, popisu a analýzy dat charakterizující životní prostředí společně s daty medicínami. Zároveň byl také v práci podrobně popsán postup prezentace těchto dat a výsledků veřejnosti.

První část práce se věnuje **získání, přípravě a popisu dat**. V České republice je velké množství dat, které z různých pohledů popisují životní prostředí. Bohužel tato data nejsou v jedné databázi, ke které by měly jednotlivé instituce přístup a postupně data doplňovaly. Z toho důvodu je nejprve nutné jednotlivé organizace oslovit a data od nich získat. Zároveň podrobnost dat je velmi rozdílná, někdy se jedná o data bodová, jindy o data polygonová. Například data od VÚKOZ jsou primárně tvořena pro prezentaci České republiky ve srovnání s Evropou. Nikdo zde nepředpokládal, že by byly využity pro jeden okres. Z toho důvodu jsou data méně podrobná, ale v ČR nám není známa podrobnější databáze, která by se zaměřovala na území okresu. Výrazným posunem vpřed v této práci by bylo získání dat, která se podrobněji zaměřují na konkrétní území.

Data od ČGS byla nejvhodnějšími získanými daty v této práci, jelikož zde byly přesně zaznamenány zlomové struktury a radiometrie v daném místě. Z toho důvodu jsou také data v aplikaci prezentována samostatně, jelikož lze díky nim názorně a dostatečně popsat daný jev. Přesto by i zde mohlo dojít ke zlepšení. Vhodné by bylo, při navázání na tuto práci získat izotopická data, která mohou hodně vypovědět o životním prostředí v daném místě. Další možností by bylo v dané oblasti udělat sběr horninových a půdních vzorků. Celkové zaměření na geologii by mohlo přinést nové poznatky ohledně vlivu podloží na člověka.

Data získaná od ČHMÚ byla dvojího typu. První datová sada, která přesně znázorňuje dané časové období, pochází z měřících stanic ČHMÚ, které jsou umístěny vždy ve městech. Z toho důvodu je rozmístění nepravidelné, ale je dostačující, jelikož pacienti bývají nejvíce koncentrováni ve městech. Datová sada, která rovnoměrně pokrývá celé území, je bohužel za období o pět let později, proto bylo nutné při její interpretaci brát tento fakt v potaz.

V neposlední řadě zde byla využita data z rozborů pitné vody, která zaznamenávala dané časové období a zároveň obsáhla většinu pacientů. V této datové sadě došlo k velkému rozšíření oproti předchozím pracím (Grundloch 2009, Stramová 2017) a z výsledků analýzy O2PLS je patrné, že toto rozšíření bylo vhodné.

Všechny tyto datové sady popisovaly z určitých úhlů životní prostředí. Byla snaha, aby v této práci bylo prostředí popsáno co nejkomplexněji a nechyběl žádný podstatný parametr. Samozřejmě by bylo možné dále rozšířit datovou základnu, ale jednalo by se

už o rozsáhlou analýzu, která by přesahovala rámec diplomové práce, jak časovou náročností na získání dat, tak jejich následným zpracováním a interpretací pro veřejnost.

Endokrinologický ústav, který poskytl medicínální data, má k dispozici čtyři datové sady popisující rozdílná území v ČR (Příbramsko, Klatovsko, Novobystřicko, Žďárské vrchy). Zpracování těchto datových sad je poměrně časově náročné, jelikož se jedná o data v tištěné podobě a jejich elektronická verze není k dispozici. Diplomová práce představila postup zpracování těchto dat. V případném navázání na tuto práci by bylo vhodné se zaměřit na zbylé datové sady. Jelikož data, poskytnutá Endokrinologickým ústavem, jsou velmi podrobná a zahrnují velké množství informací. Data byla získávána pomocí papírových dotazníků, které museli vždy pacienti vyplnit společně s odbornými lékaři. V dnešní době již existují lepší a rychlejší způsoby jak tato data získat. Jednou z možností je dotazník vytvořený v programu Survey123 od ArcGIS. Pomocí tohoto programu je možné vytvářet dotazníky o různé délce, s různými variantami odpovědí a jeho výhodou je, že přímo zaznamenává bydliště pacienta, které by pak už nemusel nikdo složitě dohledávat. Rovnou by vznikala bodová vrstva s informacemi o zdraví pacienta a o jeho bydliště. Případně by pak bylo možné informace rozšířit o místo pracoviště u dospělých či místo navštěvované školy u studentů. Takto by mohly postupně vzniknout časové řady dat, které by rozhodně přinesly velké množství zjištění.

V práci byla představena metodika, jak takto rozdílná data (bodová, liniová, polygonová; geologie, půda, ovzduší, voda, medicínální) zpracovat a připravit pro vstup do jednotlivých analýz, i pro prezentaci veřejnosti. Návaznost na tento postup zpracování měly **prostorové analýzy**. Ty již v určité míře byly provedeny v práci Stramové (2017) a jelikož jejich kvalita byla dostačující, nebylo nutné analýzy provádět znovu. Výsledkem bylo, že pacienti jsou shlukováni v obcích a jejich výskyt je vázaný na výskyt obyvatel. Tento výsledek byl potvrzen v několika analýzách, a tak již není potřeba tento aspekt dále zkoumat. Vznikl zde tedy prostor pro posouzení medicínálních dat z pohledu variogramu. Nejprve byly zjištěny základní statistické údaje jednotlivých datových sad a poté byl postupnými kroky zkoumán trend v rozložení dat. Výsledkem byl vždy variogram, který ale ve většině případů nepřinesl uspokojivé výsledky a nedošlo tedy k dostatečnému potvrzení prostorové závislosti (nezávislosti).

Dalším krokem práce byla analýza **O2PLS/OPLS**, která hodnotí korelaci vytvořené datové sady mezi sebou a dívá se na medicínální data z odlišného pohledu než prostorové analýzy. Celkově byla provedena analýza O2PLS/OPLS sedmkrát. Jednou byla aplikována pro všechny parametry lidského organismu a výsledkem byla komponenta, která popisovala míru vlivu relevantních prediktorů na alergie u člověka. Ostatní analýzy byly konkrétně zaměřeny

na parametry funkce štítné žlázy. Na tyto parametry byl dohromady identifikován vliv 10 ukazatelů životního prostředí.

Metoda O2PLS/OPLS byla aplikována vždy na celý soubor dat, který byl k dispozici. Při prvním použití byly označeny za závislé proměnné  $Y$ , byly zvoleny veškeré lékařské parametry a jako nezávisle proměnné  $X$  byly zvoleny ukazatele životního prostředí a pohlaví, věk, bmi a jodur. Výsledkem této analýzy byla pouze jedna komponenta o závislé proměnné alergie, tedy ukazatel říkající, jestli pacient trpí nějakou alergií nebo ne. Zde byla prokázána souvislost mnoha parametrů životního prostředí. Vliv těchto parametrů měl pozitivní i negativní korelaci. Zároveň u mnoha z nich byly prokázány alergické reakce lidského organismu na kontakt s těmito látkami. Jednalo se o Se v pitné vodě, Bi v mechu, BZn, Bap,  $SO_2$ , Cd,  $PM_{2,5}$ , Pb a Ni v imisích.

Následně už byly do analýzy OPLS vybírány parametry funkce štítné žlázy. Tyto parametry byly vždy testovány jednotlivě, nevybrané parametry byly zahrnuty do nezávisle proměnných  $X$ . Prvním testovaným parametrem byl VOLGUT neboli objem štítné žlázy. U tohoto parametru je známo, že štítná žláza zvětšuje svůj objem s věkem, a to se také analýzou potvrdilo. Z ukazatelů životního prostředí byl identifikován jako relevantní parametr mangan v pitné vodě. Jednalo se zde o pozitivní korelaci, tedy že větší množství Mn v pitné vodě koreluje s větším objemem štítné žlázy. Zvětšená štítná žláza je určitým projevem poruchy její funkce. Z výsledků je patrné, že přijímání Mn v pitné vodě může podporovat její růst.

Mangan v pitné vodě byl zaznamenán ještě při jedné analýze OPLS parametrů funkce štítné žlázy, a to u TSH. Zde byla ale identifikována negativní korelace, tedy čím vyšší koncentrace Mn v pitné vodě, tím nižší hladina TSH v organismu. U TSH je třeba, aby hladina byla vždy v normě a nebyla ani nízká ani vysoká. Nabízí se tedy možnost regulace této hladiny za pomoci předem určených dávek manganu. Pro TSH byl z ukazatelů životního prostředí významný právě mangan v pitné vodě a poté také koncentrace Mo v mechu,  $PM_{10}$  a As v ovzduší, Pl a B v pitné vodě. Právě  $PM_{10}$  a As v ovzduší byly identifikovány jako statisticky významnější prvky, a to na hladině pravděpodobnosti 99 %. U arsenu v ovzduší byla zjištěna také negativní korelace, což znamená, že větší množství As v ovzduší snižuje hladinu TSH v organismu. Negativní korelace nastala i u Mo v mechu a boru v pitné vodě. V případě  $PM_{10}$  a pesticidních látek se jednalo o pozitivní korelaci.

Dále byla analýza OPLS spuštěna s nastavením závisle proměnné FT3. Zde byl vybrán pouze jeden ukazatel životního prostředí jako relevantní prediktor, a to průměrná koncentrace Pb ve vodě v místě bydliště a školy. Zde byla vysoká statistická jistota na hladině pravděpodobnosti 99 % a vysoký byl i regresní koeficient. Analýza byla také aplikována na FT4,

kde byl rovněž vybrán jako relevantní prediktor pouze jeden parametr, a to chrom v pitné vodě, který měl negativní korelaci s FT4. Stejně jako u hormonu TSH i zde u FT3 a FT4 je důležité, aby hladina těchto hormonů byla v normě a nebyla příliš nízká nebo vysoká. Z tohoto důvodu záleží na velikosti dávek nebo koncentrace jednotlivých relevantních prediktorů. Při vhodném dávkování mohou být tyto parametry pro lidský organismus prospěšné, ale při vysokých dávkách by se mohlo jednat o prvky se silně negativním vlivem.

Přítomnost chromu ve vodě se projevila jako relevantní prediktor při analýze OPLS glykoproteinu TGL. Rovněž zde byla zjištěna negativní korelace, tedy čím vyšší koncentrace Cr v pitné vodě, tím nižší hladina TGL. Dalším relevantním prediktorem byl Ni v pitné vodě. U glykoproteinu TGL je nebezpečná už pouze jeho zvýšená hladina a na základě výsledků analýzy by bylo možné se domnívat, že Cr nebo Ni by mohly sloužit ke snižování hladiny TGL.

Posledním zkoumaným parametrem byl reflex Achillovy šlachy u pacientů. U tohoto parametru byl jako relevantní prediktor označen pouze jeden ukazatel životního prostředí, a to koncentrace Fe v pitné vodě. Jednalo se o negativní korelace. Pro lidské tělo je železo důležitým prvkem, který pro své fungování potřebuje. To je právě patrné i z tohoto výsledku. Při vyšších příjmech železa se zrychluje právě reflex Achillovy šlachy, který značí správné fungování štítné žlázy.

Hlavním cílem diplomové práce bylo přehledně veřejnosti prezentovat stav životního prostředí v jejich místě bydliště a zároveň informovat o možném vlivu těchto faktorů na jejich zdraví. Stav životního prostředí byl zjištěn pomocí získaných datových sad a jejich podrobnému popisu za pomoci prostorových analýz. Vliv životního prostředí byl pak odhalen pomocí analýzy O2PLS/OPLS. Poslední částí práce byla prezentace těchto výsledků. Pro prezentaci byl zvolen program StoryMap od společnosti ArcGIS. Jedná se o aplikaci, která dokáže prezentovat mapové a textové výstupy zároveň. Současně nabízí široké možnosti této prezentace, které byly prokázány v např. Atlasu Pálavy. Pro tvorbu tohoto atlasu byl použit layout s názvem Map Journal, který se nejprve jevil jako vhodný i pro prezentaci této diplomové práci. V průběhu tvorby ale došlo ke zjištění, že množství dat a analýz zde aplikovaných je tak velké, že by vše působilo nepřehledně. Proto bylo od tohoto formátu odstoupeno a byl aplikován jako hlavní layout Map Series, kam byly vkládány pomocí URL odkazu jednotlivé layouty Map Journal pro tematické odlišné kapitoly. Tento postup dal za vznik přehledné aplikaci, která dokáže prezentovat takovéto, a i mnohem větší množství dat.

## 6 Závěr

Tato diplomové práce je přínosem v několika vědních oborech z důvodu multidisciplinárního přístupu, který zde bylo nutné aplikovat. Byly zde aplikovány statistické i geostatistické metody, zároveň bylo nutné proniknout do oborů jako je geologie, hydrologie a meteorologie, důležité také byl medicínální přístup a v neposlední řadě právě tvorba a prezentace mapových výstupů, tedy kartografie. Hlavním cílem této diplomové práce byla prezentace širokého množství (tematicky, typově) datových sad. Tato prezentace byla úspěšná díky programu StoryMap od společnosti ArcGIS. Z důvodu tematického zaměření diplomové práce se může zdát, že tato prezentace je vhodná pouze pro témata mapující životní prostředí. Ale v práci jsou efektivně prezentována i data zdravotnická, což je velkým přínosem pro tento vědní obor. Zároveň je tímto potvrzeno, že pomocí online mapových výstupů je možné představit i jevy, který se mohou zdát, že nemají prostorové určení. Pro veřejnost je tato prezentace zajímavá a přitažlivá.

Ideální představou této diplomové práce by bylo sestavit takovouto prezentaci pro celou Českou republiku, aby se každý občan mohl bez problémů podívat na informace o životním prostředí v jeho bydliště a také o vlivu tohoto prostředí na jeho zdraví. K naplnění této představy by bylo potřeba vytvořit přehlednou databázi vzorku pacientů po celé republice, zároveň získat k dispozici data, která životní prostředí charakterizují v dostatečné podrobnosti a následně tato data zpracovat. Již ze samotného popisu je jasné, že toto rozmezí by bylo nad rámec diplomové práce a z toho důvodu bylo zmenšeno měřítko práce z celé České republiky na jeden okres. Metodika zde představovaná je velmi podrobná a je možné ji využít k dosažení ideálního stavu.

Dílčí cíle diplomové práce vedli právě k vytvoření aplikace a k předkládané metodice. První dílčím cílem bylo získání prostorových dat, která v dostatečné míře popisují životní prostředí v zájmovém území a následně data připravit pro vstup do analýz a pro prezentaci veřejnosti. Tato část byla naplněna díky získání datové sady z diplomové práce Stramové (2017), která byla díky spolupráci se SZÚ rozšířena o podrobnější rozbor pitné vody a dále o další geologické podklady, kterým nebyl v předchozí práci dán takový prostor. Následně byly veškeré aspekty životního prostředí popsány tak pomocí textových výstupů tak také pomocí mapových výstupů. V návaznosti na popis životního prostředí, bylo potřeba popsat chování zdravotních parametrů pacientů v prostoru nebo spíše jejich prostorové uspořádání. Tento cíl byl naplněn díky metodám prostorových analýz a také díky tvorbě a interpretaci variogramu. Významnou částí práce byla aplikovaná metoda O2PLS/OPLS, jejímž výsledkem byl regresní model pro jednotlivé



parametry funkce štítné žlázy, který v rozdílné míře objasnil souvislost parametrů funkce štítné žlázy s parametry životního prostředí.

Celý tento postup vedle ke vzniku aplikace, jež je k dispozici po obhajobě práce na tomto odkazu:

**<https://cuni.maps.arcgis.com/apps/MapSeries/index.html?appid=82411bd3e15a4fd79fb3f24bf82d6974>**.

Bylo by škoda, kdyby vytvořením webové aplikace tato diplomová práce skončila, jelikož toto téma ukrývá velký potenciál. V Endokrinologickém ústavu v Praze jsou po domluvě k dispozici medicínální data za čtyři území České republiky, které by bylo možné pomocí předkládané metodiky zpracovat. Po zpracování takového množství dat, by bylo možné dále data rozšiřovat díky vytvoření dotazníku v programu Survey123, který je online a je tedy následně snadno šiřitelný mezi lékaře. Zároveň v ČR existuje velké množství institucí i jedinců, které na základě určitého parametru zkoumají životní prostředí. Spoluprací s těmito subjekty by mohlo dojít k rozšíření datové základny popisující životní prostředí, která by pak byla využitelná i pro další vědní disciplíny. Posledním možným postupem by bylo pokračování v aplikaci geostatistických metod. V dnešní době již existují metody, které na základě vhodných podkladových dat dokáží identifikovat místa se zvýšeným rizikem určitého problému. To by posléze mohlo vést k efektivnějšímu řešení problému, ale to je již práce pro velký tým lidí, nikoli pro analýzu v rámci jedné diplomové práce.

## 7 Literatura

ARCGIS PRO (2018): Spatial Statistics toolbox, <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/an-overview-of-the-spatial-statistics-toolbox.htm> (cit. 10. 4. 2018).

ARCGIS (2018): What is ArcGIS Online?, <https://doc.arcgis.com/en/arcgis-online/reference/what-is-agol.htm> (cit. 31. 5. 2018).

ARNIKA (2014): Látky znečišťující ovzduší, <http://arnika.org/latky-znecistujici-ovzdusi> (cit. 22. 4. 2018).

ARNIKA (2014): Rtůť a zdraví, <http://arnika.org/rtut-a-zdravi> (cit. 11. 4. 2018).

ASHIZAWA, K. a kol. (1997): Prevalence of Goiter and Urinary Iodine Excretion Levels in Children Around Chernobyl 1. The Journal of Clinical Endocrinology & Metabolism, 82, 10, 3430-3433.

BARANOVIČOVÁ, Z. (2015): Scintigrafie štítné žlázy. Brno, 2015. Bakalářská práce. LF MU. Vedoucí práce Igor Černý.

BBC (2014): John Snow, [http://www.bbc.co.uk/history/historic\\_figures/snow\\_john.shtml](http://www.bbc.co.uk/history/historic_figures/snow_john.shtml) (cit. 5. 6. 2018).

BENCKO, V. a kol (2003): Statistické metody v epidemiologii. Praha, Karolinum. ISBN 80-246-0763-8.

BLAŽKOVÁ, M. (Nedatováno): Geologické faktory, důležitá složka životního prostředí. Universita J. E. Purkyně Ústí nad Labem.

BLAŽKOVÁ, M. (1996): Geologie a životní prostředí. Vysoká škola báňská- Technická univerzita Ostrava.

BØLVIKEN, B., NILSEN, R., UKKELBERG, Å.(1997): A new method for spatially moving correlation analysis in geomedicine. Environmental Geochemistry and Health, 143-153.

BRODY, G. a kol. (2004): Breast cancer risk and historical exposure to pesticides from wide-area applications assessed with GIS. Environmental Health Perspectives, 112, 8, 889.

CAHA, J. (2013): Sodík – minerální látky. In: Aktin, <https://aktin.cz/2061-sodik-mineralni-latky> (cit. 11. 4. 2018).

CESKEDALNICE.CZ (2016): Mapa české dálniční sítě, <http://www.ceskedalnice.cz/dalnicni-sit/> (cit. 25. 4. 2018).

CLARKE, B., FOKOUE, E., ZHANG, H. (2009): Principles and theory for data mining and machine learning. Springer Science & Business Media.

ČESKÝ STATISTICKÝ ÚŘAD (2017): Charakteristiky okresu Příbram, [https://www.czso.cz/csu/xs/charakteristika\\_okresu\\_pribram](https://www.czso.cz/csu/xs/charakteristika_okresu_pribram) (cit. 24. 4. 2018).

ČGS (Nedatováno): Radonové riziko, <http://www.geology.cz/extranet/vav/geologicka-rizika/radon> (cit. 30. 5. 2018).

ČHMÚ (2018): Vítr, <http://portal.chmi.cz/aktualni-situace/aktualni-stav-pocasi/ceska-republika/stanice/profesionalni-stanice/tabulky/vitr> (cit. 22. 4. 2018).

ČÍSTÉ NEBE (2018): Benzo[a]pyren, <https://www.irz.cz/node/20> (cit. 29. 5. 2018).

DIAMO (2018): DIAMO, státní podnik, odštěpný závod Správa uranových ložisek, <https://www.diamo.cz/cs/sul> (cit. 24. 4. 2018).

DRÁBOVÁ, D. a kol. (2001): Patnáct let od havárie Černobylu: Důsledky a poučení. Praha, Státní úřad pro jadernou bezpečnost.

ELLIOTT, P. a kol. (2001): Risk of adverse birth outcomes in populations living near landfill sites. *Bmj.*, 323, 7309, 363-368. ISSN 0959-8138.

ESRI (2018): What is ArcGIS?, <https://www.esri.com/en-us/about/about-esri#what-we-do> (cit. 31. 5. 2018).

GÖRNER, T. (2005): Využití GIS při posuzování vlivu přírodní a antropogenní radioaktivity na vybrané lékařské diagnózy na Klatovsku. Praha. Vedoucí práce Jiří Němeček.

GRUNDLOCH, J. (2009): Geologická a environmentální interpretace dat tyreologické epidemiologie pomocí geoinformačních systémů: srovnávací lokality Krucemburk, Budišov, Radostín nad Oslavou a Dolní Rožínka. Diplomová práce, Ústav geochemie, mineralogie a nerostných zdrojů, PŘF UK, Praha. Vedoucí práce Karel Martínek.

HAVEL, M., VÁLEK, P. (2014): Arsen. In: Arnika, <http://arnika.org/arsen> (cit. 10. 4. 2018).

HENKE, J., PETROPOULOS, G. (2013): A GIS-based exploration of the relationships between human health, social deprivation and ecosystem services: The case of Wales, UK. *Applied Geography*, 45, 77-88.

HILL, C. a kol. (1986): Iodine-131 in human thyroids in Britain following Chernobyl. *Nature*, 321, 6071, 655-656.

HODRMENT, J. (2011): Jsou odvaly uranových šachet ekologickou zátěží?, <http://www.pribramsko.eu/detail03.php?ID=3262> (cit. 24. 4. 2018).

HORÁK, J. a kol. (2000): Prostorové analýza nezaměstnanosti, [http://gisak.vsb.cz/GACR\\_PAN/Brozura/PAN.html](http://gisak.vsb.cz/GACR_PAN/Brozura/PAN.html) (cit. 22. 5. 2018).

HORÁK, J. (2002): Prostorové analýzy dat. Institut geoinformatiky, Hornickogeologická fakulta, VŠB-TU Ostrava, <http://gisak.vsb.cz/pad/index.htm> (cit. 22. 5. 2018).

CHLUPÁČ, I. (2011): Geologická minulost České republiky. Praha: Academia. Neživá příroda. ISBN 978-80-200-1961-5.

IHE (1987): Zpráva o radiační situaci na území ČSSR po havárii jaderné elektrárny Černobyl. B.m.: Institutu hygieny a epidemiologie.

IRZ (2018): Informace o látkách ohlašovaných do IRZ, <https://www.irz.cz/node/20> (cit. 14. 4. 2018).

JANOUR, Z., HOLPUCH, J., BRYCH, K., DITTRT, F., SEVERA, M. (2001): Výzkum znečištění ovzduší olovem v okolí Kovohutí Příbram. Ústav termomechaniky AV ČR.

JEŘÁBKOVÁ, L. (2004): Modelové hodnocení vývoje a využití okolí přírodních parků v Praze - na modelových územích okolí PP Šárka - Lysolaje a okolí PP Hostivař - Záběhlice. Praha. Vedoucí práce Luboš Matějček.

JEŽEK, J. (2016): Geostatistika a prostorová interpolace. Karlova univerzita v Praze, Nakladatelství Karolinum.

JEŽEK, V. (2007): Historie dobývání stříbra, olověných a železných rud na Příbramsku, <http://www.mining.cz/texty/brezovky/history.htm> (cit. 24. 4. 2018).

JIROUŠEK, J. (2018): Nebeské, <http://www.nebeske.cz/> (cit. 10. 6. 2018).

JMP (2018): Variable importance plot, <https://www.jmp.com/support/help/14/variable-importance-plot.shtml> (cit. 4. 6. 2018).

JOPKOVÁ, J., VÁLEK, P. (2014): Molybden. In: Arnika, <http://arnika.org/molybden> (cit. 10. 4. 2018).

JOPKOVÁ, J., VÁLEK, P. (2014): Selen. In: Arnika, <http://arnika.org/selen> (cit. 10. 4. 2018).

KACHLÍK, V. (2003): Geologický vývoj území České republiky. Ústav geologie a paleontologie, PŘF UK a Správa úložišť radioaktivních odpadů (SÚRAO), Praha, 1-64.

KARLOVÁ, K. (2001): Modelové hodnocení hlukové zátěže obyvatelstva v okolí nádraží Praha - Vršovice s využitím GIS. Praha.

KLEGER, V., VÁLEK, P. (2014): Nikl. In: Arnika, <http://arnika.org/nikl> (cit. 9. 4. 2018).

KNECHTSBERGER, P. (Nedatováno): Radiace, nemoc a traumata z ozáření. Radiace, [http://www.darius.cz/archeus/NT\\_radia.html](http://www.darius.cz/archeus/NT_radia.html) (cit. 23. 3. 2016).

KOLÁŘ, J. (2003): Geografické informační systémy 10. Vyd. 2. přeprac. Praha: Vydavatelství ČVUT. ISBN 80-01-02687-6.

KOVOHUTĚ PŘÍBRAM (2018): Historie, <http://www.kovopb.cz/informace-o-spolecnosti/historie/> (cit. 24. 4. 2018).

KOŽÍŠEK, F. (2009): Má Česká republika výjimku z obsahu olova v pitné vodě? In: tzinfo, <https://voda.tzb-info.cz/vlastnosti-a-zdroje-vody/6109-ma-ceska-republika-vyjimku-z-obsahu-olova-v-pitne-vode> (cit. 5. 5. 2018).

KURTIO, P. a kol. (2013): Fallout from the Chernobyl accident and overall cancer incidence in Finland. *Cancer epidemiology*, 37, 585-592.

LIV (2017): Sodík a jeho vliv na organismus. In: Zdravé zdraví, <https://zdravezdravi.cz/vitaminy/sodik-a-jeho-vliv-na-organismus> (cit. 11. 4. 2018).

MADSEN, R., LUNDSTEDT, T., TRYGG, J. (2010): Chemometrics in metabolomics—a review in human disease diagnosis. *Analytica chimica acta*, 659, 23-33.

MALÁ, B. (2013): Klasifikace komponent získaných z dat funkční magnetické rezonance pomocí analýzy nezávislých komponent. Diplomová práce, Přírodovědecká fakulta Masarykova univerzita, Brno. Vedoucí práce Daniel Schwarz.

MELOUN, M. a kol. (2000): Transformation in the PC-aided biochemical data analysis. *Clinical chemistry and laboratory medicine*, 38, 553-559.

MELOUN, M., MILITKÝ, J., HILL, M. (2005): Počítačová analýza vícerozměrných dat v příkladech. Praha: Academia. ISBN 80-200-1335-0.

MODRÁČKOVÁ, J. (1999): Využití GIS při zpracování dat z povodí Berounky. Praha.

MUCHA, J. (Nedatováno): Vlastnosti a účinky radonu na lidský organismus, <http://www.radon-servis.cz/Radon/3-Vlastnosti-a-ucinky-radonu-na-lidsky-organismus> (cit. 4. 6. 2018).

NETRDOVÁ, P. (2018): Prostorová analýza dat, [https://dl2.cuni.cz/pluginfile.php/12672/mod\\_resource/content/4/Prednaska\\_SADII\\_2016\\_Pr\\_ostorova\\_analyza.pdf](https://dl2.cuni.cz/pluginfile.php/12672/mod_resource/content/4/Prednaska_SADII_2016_Pr_ostorova_analyza.pdf) (cit. 29. 3. 2018).

NOVÁKOVÁ, M. (2018): Použití dat pro diplomovou práci [email] (cit. 12. 4. 2018).

Periodická tabulka (2017): Periodická soustava prvků, <http://www.prvky.com> (cit. 9. 4. 2018).

PETERKA, M., PETERKOVÁ, R., LIKOVSKÝ, Z. (2004): Chernobyl: prenatal loss of four hundred male fetuses in the Czech Republic. *Reproductive Toxicology*, 18, 1, 75-79.

PETRLÍK, J., VÁLEK, P. (2014): Mangan. In: Arnika, <http://arnika.org/mangan> (cit. 9. 4. 2018).

PIKNER, R. (2005): Thyreoglobulin, <http://ciselniky.dasta.mzcr.cz/cd/hypertext/MZACB.htm> (cit. 5. 5. 2018).

POULSTRUP, A., HANSEN, H. L. (2004): Use of GIS and exposure modeling as tools in a study of cancer incidence in a population exposed to airborne dioxin. Environmental health perspectives, 112, 9, 1032.

RADONOVÝ PROGRAM ČR (2016): Radon v ČR, <https://www.radonovyprogram.cz/radon-v-cr/> (cit. 30. 5. 2018).

REINERS, CH., BIKO, J., DEMIDCHIK, E., DROZD, V. (2000): Thyroid cancer after exposure to ionising irradiation: histology, staging and clinical data. The Thyroid and Environment, Budapest 2000.

SCS (2018): Pojem statistika, <https://slovník-cizích-slov.abz.cz/web.php/slovo/statistika> (cit. 5. 6. 2018).

SHAHID, S. U., IQBAL, J., HASNAIN, G. (2014): Groundwater quality assessment and its correlation with gastroenteritis using GIS: a case study of Rawal Town, Rawalpindi, Pakistan. Environmental monitoring and assessment, 186, 11, 7525-7537.

SPURNÁ, P. (2008): Geograficky vážená regrese: metoda analýzy prostorové nestacionarity geografických jevů. Geografie – Sborník ČGS, 113, 2, 125 – 139.

SPURNÁ, P. (2008): Prostorová autokorelace – všudypřítomný jev při analýze prostorových dat. Sociologický časopis, 44, 4, 767 – 787.

STÁTNÍ ZDRAVOTNÍ ÚSTAV (2016): Poslání ústavu, <http://www.szu.cz/poslani-ustavu> (cit. 25. 2. 2018).

STRAKA, J. (2018): Periodická tabulka, <http://www.tabulka.cz> (cit. 9. 4. 2018).

STRAMOVÁ, A. (2017): Geoinformační analýza zdravotních dat. Diplomová práce, Katedra aplikované geoinformatiky a kartografie, PřF UK, Praha. Vedoucí práce Lukáš Brůha.

SUCHAROVÁ, J., SUCHARA, I., HOLÁ, M. (2001): Určení zátěže krajiny ČR aktuálními úrovněmi atmosférické depozice a zajištění dat pro celoevropský biomonitorovací program 2000. Projekt MŽP VaV 640/1/00.

SURO (2018): Přírodní radioaktivita a problematika radonu, <https://www.suro.cz/cz/prirodnioz/obecne-informace/zdravotni-ucinky-radonu> (cit. 30. 5. 2018).

SURO (2018): Více o zdravotních účincích RADONU, <https://www.suro.cz/cz/prirodnioz/obecne-informace/zdravotni-ucinky-radonu> (cit. 30. 5. 2018).

ŠTYCH, P. a kol. (2008): Vybrané funkce geoinformačních systémů. Praha: Česká kosmická kancelář.

TOLLINGEROVÁ, D. (1996): GIS. Geografické informační systémy. Ostrava: VŠB-Technická univerzita, Phare. ISBN 80-7078-377-X.

TOMANOVÁ, R. (2005): Vliv geologie a radioaktivního záření na výskyt dětské tyreopatie: testovaná lokalita Příbramsko. Rigorózní práce, Ústav geologie a paleontologie, PŘF UK, Praha. Vedoucí práce Karel Martínek.

TOMSA, J. (2011): Odvození normální rozdělení, <http://mujweb.cz/jtomsa/gauss.htm> (cit. 22. 3. 2018).

TONDEL, M. a kol. (2004): Increase of regional total cancer incidence in north Sweden due to the Chernobyl accident. Journal of epidemiology and community health, 58, 12, 1011-1016.

TRYGG, J., HOLMES, E., LUNDSTEDT, T. (2007): Chemometrics in metabonomics. Journal of proteome research, 6, 2, 469-479.

TRYGG, J., WOLD, S. (2003): O2-PLS, a two-block (X–Y) latent variable regression (LVR) method with an integral OSC filter. Journal of Chemometrics, 17, 1, 53-64.

TRYGG, J., WOLD, S. (2002): Orthogonal projections to latent structures (O-PLS). Journal of chemometrics, 16, 3, 119-128.

UČV (2015): Pitná voda, <https://www.ucv.cz/2016/04/05/pitna-voda/> (cit. 10. 4. 2018).

VÁLEK, P. (2014): Benzo(a)pyren. In: Arnika, <http://arnika.org/benzen> (cit. 29. 5. 2018).

VÍTEJ NA ZEMI (2018): Znečištění ovzduší benzo(a)pyrenem, <https://www.irz.cz/node/20> (cit. 29. 5. 2018).

VOŽENÍLEK, V. (1998): Geografické informační systémy I: pojetí, historie, základní komponenty. Olomouc: Vydavatelství Univerzity Palackého. ISBN 80-7067-802-X.

VÚKOZ (2016): Historie, <http://www.vukoz.cz/index.php/cs/o-nas/historie> (cit. 25. 2. 2018).

ZAMRAZIL, V. (2002): Choroby štítné žlázy - Základy diagnostiky a terapie. In: Postgraduální medicína, <https://zdravi.euro.cz/clanek/postgradualni-medicina/choroby-stitne-zlazy-zaklady-diagnostiky-a-terapie-144038> (cit. 5. 5. 2018).

ZHANG, Ch. (2006): Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. *Environmental pollution*, 142, 3, 501-511.



## Seznam příloh

Příloha 1	Doprovodné texty použité v aplikaci
Příloha 2	Seznam vybraných parametrů endokrinologických dat